ANÁLISIS MATEMÁTICO

Curso intermedio

BIBLIOTECA DE MATEMÁTICA SUPERIOR
bajo la dirección del
Dr. Emilio Lluis Riera

Traducción: **Federico Velasco Coba**
Director del Departamento
de Matemáticas
Facultad de Ciencias
Universidad Veracruzana

Revisión técnica: **Emilio Lluis Riera**
Instituto de Matemáticas
Facultad de Ciencias
Universidad Nacional Autónoma
de México

Supervisión editorial: **Federico Galván Anaya**
Catedrático de Matemáticas
Universidad Nacional Autónoma
de México

BIBLIOTECA DE MATEMÁTICA SUPERIOR

ANÁLISIS MATEMÁTICO
Curso intermedio

BIBLIOTECA DE MATEMÁTICA SUPERIOR

Volumen 2

Norman B. Haaser
Joseph P. La Salle
Joseph A. Sullivan



EDITORIAL TRILLAS



México, Argentina, España, Colombia, Puerto Rico, Venezuela

Catalogación en la fuente

Haaser, Norman B.
Análisis matemático 2 : curso intermedio. -- 2a ed. -- México : Trillas, 1990 (reimp. 1995).
v. 2 (786 p.) ; 23 cm. -- (Biblioteca de matemática superior)
Traducción de: Intermediate analysis
Bibliografía: p. 777-779
Incluye índices
ISBN 968-24-3882-9

1. Análisis matemático. I. LaSalle, Joseph P. II. Sullivan, Joseph A. III. t. IV. Ser.

LC-QA37'H3.3 D-510'H736a 219

Título de esta obra en inglés:
Intermediate Analysis

Versión autorizada en español de la
primera edición publicada en inglés por
© Blaisdell Publishing Company
A division of Ginn and Company
Waltham, Massachusetts, E. U. A.

Av. Río Churubusco 385, Col. Pedro María Anaya,
C. P. 03340, México, D. F.

División Comercial, Calz. de la Viga 1132, C. P. 09439
México, D. F. Tel. 6330995, FAX 6330870

Miembro de la Cámara Nacional de la
Industria Editorial. Reg. núm. 158

Primera edición en español, 1970 (ISBN 968-24-0142-9)
Reimpresiones, febrero y noviembre 1971, 1972,
mayo y septiembre 1973, 1974, 1975, 1976, 1977,
1979, 1980, 1982, 1983, 1985, 1986, 1987 y 1989
Segunda edición en español, 1990 (ISBN 968-24-3882-9)
Reimpresión, 1992

Segunda reimpresión, febrero 1995

Impreso en México
Printed in Mexico

Prólogo

Este libro se escribió pensando hacer de él un libro de texto para un segundo curso de matemáticas a nivel universitario. Presupone una introducción al cálculo de funciones reales de una variable real. Aunque es el segundo volumen de una serie no supone, sin embargo, que el estudiante debe haber estudiado el volumen I, *Introducción al análisis*, de los mismo autores. Pero sí suponemos que el lector ha estudiado el sistema de los números reales y está familiarizado con sus propiedades fundamentales y que también le son familiares las ideas de límite, derivada e integral.

Este nuevo volumen ofrece al estudiante otra oportunidad para aumentar su comprensión y apreciación de las ideas fundamentales del análisis. La geometría y el cálculo se extienden en dimensión con los vectores n-dimensionales. Mucho de este material debe ser ya familiar al estudiante, pero aquí aparece en un contexto más general. Le presentamos al lector numerosas extensiones y nuevas técnicas e introducimos nuevos e importantes conceptos.

Nosotros vemos el análisis no sólo como matemáticas, sino también como un instrumento de la ciencia. Según nos ha parecido posible y práctico presentamos el análisis a la luz de las matemáticas contemporáneas. La técnica es importante y necesaria tanto para el matemático como para el que usa las matemáticas, pero si algo nos ha enseñado la marcha del desarrollo científico, ello ha sido la supremacía de las ideas. Las ideas y las relaciones de las ideas hacen interesantes e inteligibles las matemáticas; la apreciación de las ideas las hace útiles.

El texto está dividido en forma que creemos natural, en cinco unidades, y es posible adaptarlo a una extensa variedad de cursos.

En los capítulos 1 y 2 se discuten el álgebra de los vectores en el espacio n-dimensional y la geometría del espacio n-dimensional con énfasis particular en el espacio de tres dimensiones. Para los estudiantes ya familiarizados con los vectores y el enfoque vectorial de la geometría bidimensional, los primeros dos capítulos le procuran un repaso de este conocimiento, al mismo tiempo que extienden sus ideas a dimensiones más altas. Para tales estudiantes, el tiempo que deben dedicar a estos capítulos puede ser muy breve. Para los que no estén familiarizados con los vectores y el enfoque vectorial de la geometría es para los que hemos elaborado estos capítulos ampliamente.

Los capítulos 3, 4 y 5 están fundamentalmente dedicados a generalizar el cálculo diferencial de funciones reales de una sola variable real para los casos donde el rango es un conjunto de vectores, donde lo es el dominio, y donde tanto el dominio como el rango lo son, respectivamente. Estos capítulos proporcionan al estudiante oportunidades adicionales de conseguir

una mejor comprensión de los conceptos de límite, continuidad y derivada, que aparecen como generalizaciones naturales a espacios de más alta dimensión.

En los capítulos 6 y 7 se generaliza el cálculo integral de funciones reales de una variable real a las funciones reales de un vector, es decir, a funciones reales de diversas variables reales. Se estudian primero las integrales dobles y se da a continuación un breve tratamiento de las integrales triples. En la sección 19 del capítulo 6; se enuncia la mayoría de los resultados de las secciones anteriores del capítulo, pero para n-dimensiones se indican cuáles serán las modificaciones que deben hacerse en las pruebas anteriores. En el capítulo 7 introducimos las funciones de conjunto (sobre familias de conjuntos) y probamos teoremas análogos a los teoremas fundamentales del cálculo. Obtenemos después la fórmula para el cambio de variables para las integrales triples; primero para cambios lineales y luego para transformaciones de clase C^1.

Los capítulos 8, 9 y 10 tratan de sucesiones, series infinitas y el tópico, íntimamente relacionado con los anteriores, de las integrales impropias (infinitas). En el capítulo 8 se presenta una extensa discusión de las sucesiones como previa al tratamiento de las series infinitas del capítulo 9. En la sección 7 del capítulo 8 se prueban algunos teoremas sobre funciones continuas de un vector (funciones de diversas variables reales). Estos resultados se han estado usando sin prueba en partes anteriores del libro, pero siempre haciendo referencia a esta sección. El capítulo 10, aunque estrictamente no depende de los capítulos 8 y 9, hace uso de la analogía entre series infinitas e integrales impropias. Aparte del estudio de las integrales impropias, en el capítulo 10 se derivan también algunas propiedades de las integrales definidas dependientes de un parámetro.

En los capítulos 11 y 12 se introducen las ecuaciones diferenciales. El capítulo 11 comienza con las ecuaciones lineales de primer orden. Sigue a esto una discusión de los sistema lineales bidimensionales con coeficientes constantes. Los resultados para las ecuaciones lineales de segundo orden con coeficientes constantes se obtienen directamente partiendo del trabajo previo sobre sistemas bidimensionales. Los métodos aquí usados pueden generalizarse fácilmente a sistemas de dimensión mayor y a ecuaciones de orden más alto. Siguen después discusiones sobre oscilaciones lineales, ecuaciones exactas y curvas integrales. El capítulo 12 comienza con la discusión de un teorema de punto fijo y nos lleva al teorema de existencia y unicidad de Picard-Lindelöf para las ecuaciones diferenciales. La definición de funciones por ecuaciones diferenciales y el estudio de las propiedades de tales funciones se ilustra a continuación. Concluye el capítulo con una introducción al estudio de un tópico íntimamente relacionado con el precedente; el de las series y las aproximaciones de Fourier.

Estamos profundamente agradecidos a los profesores René DeVogelaere,

Lester Lange y Richard Otter por sus muchos y útiles comentarios y sugerencias al comienzo de este trabajo; al profesor Richard Bishop que leyó cuidadosamente todo el manuscrito en un primer borrador y nos dio una lista de errores y comentarios; y al profesor Harley Flanders que nos hizo numerosas sugerencias después de revisar el manuscrito. Apreciamos también en todo su valor las oportunidades dadas por la Universidad de Notre Dame y el Colegio de Boston al permitirnos experimentar en nuestras cátedras. Aunque J.P. LaSalle no participó en estos experimentos, tomó parte en el planeamiento y la primera redacción de este volumen y preparó el manuscrito de los capítulos 11 y 12. Una palabra final de gratitud a nuestros estudiantes que, desde 1957, han estado usando este libro en ediciones preliminares y nos han permitido calibrar la conveniencia de nuestra presentación.

N. B. HAASER
J. A. SULLIVAN

Índice general

Capítulo 4
FUNCIONES REALES DE UN VECTOR **163**

Capítulo 5
FUNCIONES VECTORIALES DE UN VECTOR **249**

Capítulo 6
INTEGRALES MÚLTIPLES **311**

Índice de símbolos

		Página
V_n	espacio vectorial n-dimensional	16
$\mathbf{x}$	un vector en V_n	16
R	el sistema de los números reales	16
$\lvert\mathbf{a}\rvert$	la longitud de **a**	26
$\mathbf{a}\cdot\mathbf{b}$	producto escalar	29
$V_n(F)$	espacio vectorial n-dimensional sobre el campo F	37
R^3	espacio euclidiano tridimensional	44
$\mathbf{a}\times\mathbf{b}$	producto vectorial	54
$[\mathbf{abc}]$	triple producto escalar	59
$\Rightarrow$	implica	74
$\Leftrightarrow$	si y sólo si	75
$[a, b]$	intervalo cerrado	87 y sigtes.
$\langle a, b\rangle$	intervalo abierto	87 y sigtes.
I	función identidad	99
$\mathscr{S}(\mathbf{c}; r)$	vecindad de **c** de radio r	102
$\mathscr{S}'(\mathbf{c}; r)$	vecindad reducida de **c** de radio r	102
$\mathbf{f}'$	derivada de **f**	115
$f\circ g$	f composición g	125
$\mathbf{T}(t)$	vector tangente unitario	143
$\mathbf{N}(t)$	vector normal principal unitario	144
$\mathbf{B}(t)$	vector binormal unitario	146
$\kappa(t)$	curvatura	150
$\rho(t)$	radio de curvatura	150
$\tau(t)$	torsión	152
$\mathscr{C}\mathscr{E}$	complemento de un conjunto $\mathscr{E}$	164
$\mathscr{E}_i$	interior de $\mathscr{E}$	164
$\mathscr{E}_b$	frontera de $\mathscr{E}$	164
$\mathscr{E}_e$	exterior de $\mathscr{E}$	164
$\bar{\mathscr{E}}$	cerradura de $\mathscr{E}$	166
I_k	función proyección	171
$df(\mathbf{x}; \mathbf{h})$	diferencial de f	189
$\mathbf{D}f$	derivada de f	189
$D_{\mathbf{u}}f$	derivada direccional	196
$D_k f$	derivada parcial	202
$\dfrac{\partial z}{\partial x}$	derivada parcial	205
∇f	gradiente de f	207
$\dfrac{\partial^2 w}{\partial x_j \partial x_k}$	derivada parcial de segundo orden	213





Álgebra vectorial

1. INTRODUCCIÓN

En los asuntos cotidianos de nuestras vidas y aún más en la ciencia es útil e incluso a veces esencial describir con números objetos, eventos y fenómenos. En algunos casos un solo número nos basta. Por ejemplo la distancia entre Nueva York y Londres puede darse por un solo número. Sin embargo, la localización de una ciudad sobre la Tierra requiere dos números, y la localización de un objeto en el espacio requiere tres. En la física, magnitudes tales como fuerza, desplazamiento, velocidad, aceleración y momento pueden especificarse por tres números. Hay muchos ejemplos en que, para describir una situación física, se necesitan más de tres números. Por ejemplo, para localizar una partícula en el espacio y el tiempo se necesitan cuatro números. La descripción del estado del mercado de valores

o el estado de un circuito eléctrico pueden fácilmente exigir el empleo de miles de números.

Todos los ejemplos anteriores en que una colección de números especifica una magnitud física o una situación física, química, económica o social, son ejemplos de vectores. Como en el caso de los números reales, sobre los vectores podemos definir operaciones. El conjunto de todos los vectores especificados por un número fijo de números reales con ciertas operaciones básicas definidas sobre estos vectores, se llama espacio vectorial y el estudio de estas operaciones sobre los vectores se llama álgebra vectorial o lineal. El número de números reales necesarios para especificar los vectores en un espacio vectorial es la dimensión del espacio. Así, un vector en el espacio de dimensión cuatro es una cuaterna de números reales y, en general, un vector en un espacio n-dimensional es una n-ada de números reales.

Como las operaciones que deben definirse sobre los vectores y sus propiedades básicas no dependen de cuál sea la dimensión del espacio, comenzamos con el estudio de las propiedades algebraicas de los espacios vectoriales n-dimensionales. En el próximo capítulo, usaremos el álgebra de los vectores en el espacio tridimensional en el estudio de la geometría analítica sólida. En los capítulos que siguen nos ocuparemos principalmente de los vectores en los espacios de dos y tres dimensiones.

2. VECTORES

2.1 Definición. *El **espacio vectorial** **n-dimensional** V_n es el conjunto de todas las n-adas de números reales, a las que denotaremos por* $\mathbf{x} = (x_1, \ldots, x_n)$, $x_i \in \mathrm{R}$,[1] $(i = 1, \ldots, n)$ *y llamaremos vectores, donde las relaciones de igualdad y las operaciones de adición y de multiplicación por un número real se definen como sigue :*

2.2 Igualdad de vectores. *Si* $\mathbf{x} = (x_1, \ldots, x_n)$ *y* $\mathbf{y} = (y_1, \ldots, y_n)$ *son vectores en* V_n, *entonces*

$$\mathbf{x} = \mathbf{y} \quad si \quad x_i = y_i \quad para\ todo \quad i = 1, \ldots, n.$$

2.3 Adición de vectores. *Si* $\mathbf{x} = (x_1, \ldots, x_n)$ *y* $\mathbf{y} = (y_1, \ldots, y_n)$ *son vectores en* V_n, *entonces*

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n).$$

2.4 Multiplicación de un vector por un número real. *Si* $\mathbf{x} = (x_1, \ldots, x_n)$ *es un vector en* V_n *y r es un número real, entonces*

$$r\mathbf{x} = (rx_1, \ldots, rx_n).$$

[1] En todo este volumen representaremos al conjunto de los numeros reales por R. La notación $x_i \in \mathrm{R}$ se lee "x_i pertenece a R", "x_i es un elemento de R", o en forma más breve "x_i está en R".

En este libro los vectores se representan con letras negritas, $\mathbf{x}$. Los vectores suelen también representarse por símbolos tales como: $\vec{x}$, $\bar{x}$, $\underline{x}$, o $\underset{\sim}{x}$.

El número x_i se llama *i*-**ésimo componente** del vector

$$\mathbf{x} = (x_1, \ldots, x_n).$$

La relación de igualdad y la operación de adición y multiplicación por un número real pueden expresarse verbalmente como sigue:

2.2′ *Dos vectores de V_n son iguales si sus componentes correspondientes son iguales.* Por ejemplo, el vector $\mathbf{x} = (4, 0, -8, 4, 7)$ no es igual al vector $\mathbf{y} = (4, 0, -8, 7, 4)$.

2.3′ *La suma de dos vectores es el vector obtenido sumando los componentes correspondientes.*
Por ejemplo, si $\mathbf{x} = (3, 16, -2, 6, 10)$ y $\mathbf{y} = (34, -16, -4, 5, 27)$, entonces

$$\begin{aligned}\mathbf{x}+\mathbf{y} &= (3+34, 16+(-16), -2+(-4), 6+5, 10+27)\\ &= (37, 0, -6, 11, 37).\end{aligned}$$

2.4′ *El producto de un número real r por un vector $\mathbf{x}$ es el vector que se obtiene al multiplicar cada componente de $\mathbf{x}$ por el número real r.*
Por ejemplo,

$$\tfrac{1}{2}(-1, 0, 8) = (\tfrac{1}{2}(-1), \tfrac{1}{2}(0), \tfrac{1}{2}(8)) = (-\tfrac{1}{2}, 0, 4).$$

Como las operaciones de adición de vectores y multiplicación de un vector por un número real son operaciones sobre los componentes de los vectores y los componentes son números reales, las propiedades algebraicas de los números reales inducen ciertas propiedades algebraicas correspondientes en V_n.

2.5 Ejemplo. Establézcase la ley conmutativa para la adición de vectores: $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$ para todos los vectores $\mathbf{x}, \mathbf{y} \in V_n$.

SOLUCIÓN. Sean $\mathbf{u} = \mathbf{x}+\mathbf{y}$ y $\mathbf{v} = \mathbf{y}+\mathbf{x}$. Entonces, según 2.3

$$u_i = x_i + y_i, \quad v_i = y_i + x_i \quad (i = 1, \ldots, n).$$

Según la ley conmutativa para la adición de los números reales

$$u_i = v_i \quad (i = 1, \ldots, n).$$

Por tanto, de acuerdo con 2.2, $\mathbf{u} = \mathbf{y}$, es decir, $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$.

2.6 Ejemplo. Demuéstrese que: $\mathbf{0} = (0, \ldots, 0)$ es el único vector con la propiedad de que $\mathbf{x}+\mathbf{0} = \mathbf{x}$ para toda $\mathbf{x} \in V_n$.

SOLUCIÓN. Es claro que $\mathbf{x}+\mathbf{0} = \mathbf{x}$ para toda $\mathbf{x} \in V_n$. Supongamos que $\mathbf{0}'$ es

otro vector con la misma propiedad: $\mathbf{y}+\mathbf{0}' = \mathbf{y}$ para todo $\mathbf{y} \in V_n$. Entonces, tomando $\mathbf{x} = \mathbf{0}'$ y $\mathbf{y} = \mathbf{0}$, de acuerdo con el ejemplo 2.5 obtenemos

$$\mathbf{0}' = \mathbf{0}'+\mathbf{0} = \mathbf{0}+\mathbf{0}' = \mathbf{0}.$$

De donde $\mathbf{0}$ es el único vector con esta propiedad.

2.7 Ejemplo. Establézcase la siguiente ley distributiva para vectores:

$$r(\mathbf{x}+\mathbf{y}) = r\mathbf{x}+r\mathbf{y} \text{ para cualesquiera } \mathbf{x},\ \mathbf{y} \in V_n \text{ y todo } r \in \mathrm{R}.$$

SOLUCIÓN

$$\begin{aligned} r(\mathbf{x}+\mathbf{y}) &= r(x_1+y_1, \ldots, x_n+y_n) && [2.3] \\ &= (r(x_1+y_1), \ldots, r(x_n+y_n)) && [2.4] \\ &= (rx_1+ry_1, \ldots, rx_n+ry_n) && [\text{ley distributiva para R} \\ &= (rx_1, \ldots, rx_n)+(ry_1, \ldots, ry_n) && [2.3] \\ &= r\mathbf{x}+r\mathbf{y}. && [2.4] \end{aligned}$$

Y esto completa la prueba.

De modo análogo al empleado en los anteriores ejemplos, cada una de las siguientes *propiedades algebraicas fundamentales* del espacio vectorial n-dimensional V_n se pueden establecer con facilidad:

2.8 Teorema

A_1. *Para* $\mathbf{x}$ *y* $\mathbf{y}$ *de* V_n *cualesquiera*, $\mathbf{x}+\mathbf{y} \in V_n$.
A_2. *Para cualesquiera* $\mathbf{x}$ *y* $\mathbf{y}$ *en* V_n, $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$.
A_3. *Para cualesquiera* $\mathbf{x}$, $\mathbf{y}$ *y* $\mathbf{z}$ *en* V_n, $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$.
A_4. *Hay un y sólo un vector en* V_n *—denotado por* $\mathbf{0}$ *y llamado vector* cero— *con la propiedad de que*

$$\mathbf{x}+\mathbf{0} = \mathbf{x} \quad \textit{para toda} \quad \mathbf{x} \in V_n. \quad (\mathbf{0} = (0, \ldots, 0).)$$

A_5. *Para cada* $\mathbf{x} \in V_n$ *hay un vector único —denotado por* $-\mathbf{x}$*— con la propiedad de que*

$$\mathbf{x}+(-\mathbf{x}) = \mathbf{0}. \quad (-\mathbf{x} = (-1)\mathbf{x}.)$$

S_1. *Para todo* $\mathbf{x} \in V_n$ *y todo* $r \in \mathrm{R}$, $r\mathbf{x} \in V_n$.
S_2. *Para todo* $\mathbf{x} \in V_n$, $1 \cdot \mathbf{x} = \mathbf{x}$.
S_3. *Para* r *y* s *pertenecientes a* R *cualesquiera y todo* $\mathbf{x} \in V_n$, $r(s\mathbf{x}) = (rs)\mathbf{x}$.
S_4. *Para cualesquiera* $r, s \in \mathrm{R}$ *y todo* $\mathbf{x} \in V_n$, $(r+s)\mathbf{x} = r\mathbf{x}+s\mathbf{x}$.
S_5. *Para todo* $r \in \mathrm{R}$ *y cualesquiera* $\mathbf{x}$, $\mathbf{y} \in V_n$, $r(\mathbf{x}+\mathbf{y}) = r\mathbf{x}+r\mathbf{y}$.

Las propiedades A_1, A_4 y S_5 se establecieron en los ejemplos 2.5, 2.6 y 2.7. Las pruebas de las restantes propiedades son también sencillas consecuencias de las propiedades de los números reales (problema 2).

Nota. Las propiedades A_1 a A_5 del teorema 2.8 implican que el conjunto de vectores n-dimensionales es un grupo conmutativo bajo la operación de adición.

La sustracción de vectores puede definirse en términos de adición del siguiente modo.

2.9 Definición (sustracción). *Para* $\mathbf{x}$, $\mathbf{y} \in V_n$ *cualesquiera*

$$\mathbf{x}-\mathbf{y} = \mathbf{x}+(-\mathbf{y});$$

es decir, $\mathbf{x}-\mathbf{y} = (x_1-y_1, \ldots, x_n-y_n)$.

Problemas

1. Sean $\mathbf{a} = (3, -5, 4)$, $\mathbf{b} = (2, 5, 7)$, $\mathbf{c} = (0, 2, 1)$, $\mathbf{P}_0 = (0, 5, 6)$, y $\mathbf{P}_1 = (-1, -5, 2)$.

Encuéntrense:

a) $\mathbf{a}+\mathbf{b}$ b) $\mathbf{a}-\mathbf{b}$
c) $3\mathbf{a}+4\mathbf{b}-3\mathbf{c}$ *d)* $\mathbf{x}$ si $4\mathbf{x}+\mathbf{a} = 3\mathbf{b}$
e) $\mathbf{P}_0+t(\mathbf{P}_1-\mathbf{P}_0)$ *f)* $\frac{1}{2}(\mathbf{P}_0+\mathbf{P}_1)$
g) $\mathbf{P}_0+t\mathbf{a}$; $t = 0, \pm 1, \pm 2, \pm 3$
h) $\mathbf{P}_0+s\mathbf{a}+t\mathbf{b}$; $(s, t) = (0, 0), (1, 0), (0, 1), (1, 1), (-1, -1)$.

2. Pruébense las siguientes partes del teorema 2.8:

a) A_1 *b)* A_3 *c)* A_5 *d)* S_1
e) S_2 *f)* S_3 *g)* S_4.

3. Demuéstrese que: $0\mathbf{x} = \mathbf{0}$ y $r\mathbf{0} = \mathbf{0}$.

4. Demuéstrese que:

a) si $\mathbf{a}+\mathbf{b} = \mathbf{a}+\mathbf{c}$, entonces $\mathbf{b} = \mathbf{c}$
b) si $r\mathbf{x} = \mathbf{0}$, entonces $r = 0$ o $\mathbf{x} = \mathbf{0}$
c) si $r\mathbf{x} = s\mathbf{x}$ y $\mathbf{x} \neq \mathbf{0}$, entonces $r = s$

5. Demuéstrese que: Si $t \neq 0$, entonces

$$s\mathbf{a}+t\mathbf{x} = \mathbf{b}$$

tiene la única solución $\mathbf{x} = \dfrac{1}{t}(\mathbf{b}-s\mathbf{a})$.

6. Resuélvanse:

a) $2(0, 3)+8\mathbf{x} = (1, -7)$
b) $-3(1, -3, 5)+2\mathbf{x} = 5(0, -2, -1)+3\mathbf{x}$
c) $3[\mathbf{x}-(8, -3, -2, 1)] = 6(7, 0, -5, -10)$

7. En cada una de las siguientes ecuaciones determínese si hay o no números reales r que las satisfagan:

a) $(3, -2) = r(6, 4)$ *b*) $(3, -2) = r(-6, 4)$
c) $r(1, -12, 8, 13) = (3, -36, 24, 40)$
d) $r(4, 2, 0, 5) + 3(4, -2, 6, 0) = 2(6, -3, 9, 0)$
e) $2r(4, 6, -10) + 3(-2, 4, 8) = 2(-3, 6, 12) + 4r(2, 3, -5)$.

8. En cada una de las siguientes ecuaciones encuéntrense todos los números reales r y s que las satisfacen:

a) $r(3, -2) + s(6, 4) = \mathbf{0}$ *b*) $r(3, -2) + s(6, -4) = \mathbf{0}$
c) $r(8, -2, 14) + s(-12, 3, 2) = \mathbf{0}$
d) $(5, 5) = r(5, 1) + s(3, 5)$
e) $(11, 14, -2) = r(3, 0, -5) + s(1, -2, -4)$.

3. REPRESENTACIÓN GEOMÉTRICA DE LOS VECTORES

En esta sección discutiremos las ideas geométricas intuitivas que están en el fondo del álgebra vectorial y que nos guían en la construcción de nuestro modelo analítico de espacio euclidiano n-dimensional. Describiremos el modo en que los vectores en el espacio tridimensional pueden representarse por "flechas" (también se les denomina "segmentos dirigidos"). Mediante construcciones con estas flechas posteriormente dibujaremos diagramas que ilustren el álgebra vectorial. Aunque esta imagen de un vector como objeto geométrico concreto está limitada a los espacios vectoriales uni, bi, o tridimensionales, el lenguaje utilizado para los vectores de los espacios n-dimensionales se deriva de esta representación geométrica.

Escojamos en un espacio tridimensional (figura 1): 1) un punto **O**; 2) tres rectas perpendiculares entre sí X_1, X_2 y X_3 que pasen por **O**; 3) direcciones positivas sobre estas tres rectas; y 4) una unidad para medir las distancias. Un sistema como el descrito se llama "sistema cartesiano" o de "coordenadas rectangulares". *Convenimos desde ahora en que siempre limitaremos nuestras ilustraciones a sistemas de coordenadas levógiros o "de mano derecha"*. Esto significa que las direcciones positivas de las tres rectas (llamadas "ejes") han sido escogidas de tal modo que cuando el pulgar de la mano derecha apunta en la dirección positiva del eje X_1 y el dedo índice (de dicha mano) apunta en la dirección positiva del eje X_2, el dedo cordial (el de en medio) puede señalar la dirección positiva del eje X_3. Los sistemas levógiros de coordenadas pueden también describirse diciendo que la rotación en el plano $X_1 X_2$ de 90° de la semirrecta positiva del eje X_1 a la semirrecta positiva del eje X_2 es contraria a la dirección de giro de las manecillas del reloj (es "hacia la izquierda", es decir, levógira) cuando se ve desde la semirrecta positiva del eje X_3. Aunque no usaremos sistemas de

coordenadas dextrógiros en este libro, estos sistemas son de uso común en muchas ocasiones y pueden describirse reemplazando en las descripciones anteriores las palabras "mano derecha" y "contraria a la dirección de giro de las manecillas del reloj" por "mano izquierda" e "igual a la dirección del giro de las manecillas del reloj", respectivamente.

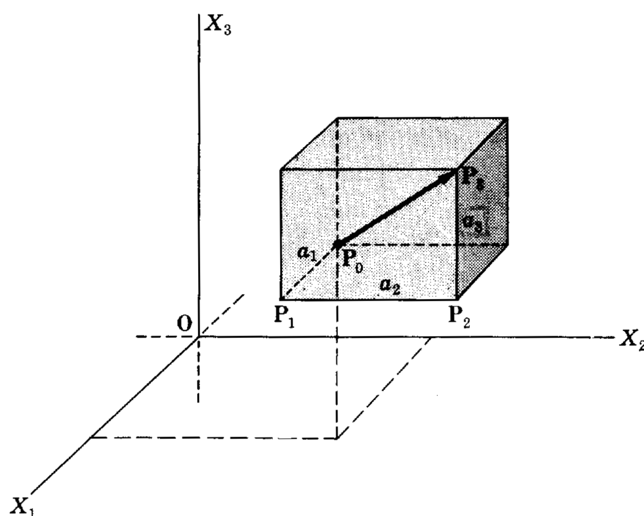


FIGURA 1

Dado un vector $\mathbf{a} = (a_1, a_2, a_3)$ en V_3, construimos una flecha que represente el vector **a** como sigue (figura 1): elegimos un punto arbitrario $\mathbf{P}_0$; nos movemos la distancia a_1 paralelamente al eje X_1 desde $\mathbf{P}_0$ y localizamos el punto $\mathbf{P}_1$ (el número a_1 es una distancia dirigida; a_1 positivo significa que debemos movernos en la dirección positiva del eje X_1 y a_1 negativo que debemos movernos en la dirección opuesta); de $\mathbf{P}_1$ nos movemos la distancia dirigida a_2 paralelamente al eje X_2 y localizamos así el punto $\mathbf{P}_2$; nos movemos de $\mathbf{P}_2$ la distancia dirigida a_3 paralelamente al eje X_3 y localizamos el punto $\mathbf{P}_3$. La flecha de $\mathbf{P}_0$ a $\mathbf{P}_3$, que también denotaremos por **a**, es una representación geométrica del vector **a**. A $\mathbf{P}_0$ se le llama punto inicial de la flecha **a** y al punto $\mathbf{P}_3$ su punto terminal. Recíprocamente, dada una flecha de $\mathbf{P}_0$ a $\mathbf{P}_3$, construyendo un paralelepípedo rectangular del que $\mathbf{P}_0$ y $\mathbf{P}_3$ sean vértices opuestos y con caras paralelas a los planos X_1X_2, X_2X_3, y X_3X_1, un vector $\mathbf{a} = (a_1, a_2, a_3)$ puede asignarse a una cualquiera de tales flechas.

Al construir la flecha que representa un vector **a**, elegimos arbitrariamente el punto inicial $\mathbf{P}_0$. Un mismo vector **a** puede estar representado por flechas diferentes. En algunas aplicaciones se establecen restricciones sobre la localización de $\mathbf{P}_0$. Por ejemplo, puede ser que se especifique cuál ha de

ser el punto inicial. Si el punto inicial es el origen **O**, el vector se llama "radio vector". En cualquier caso, el vector $\mathbf{a} = (a_1, a_2, a_3)$ determina tanto la longitud como la dirección de la flecha; dos flechas cualesquiera que

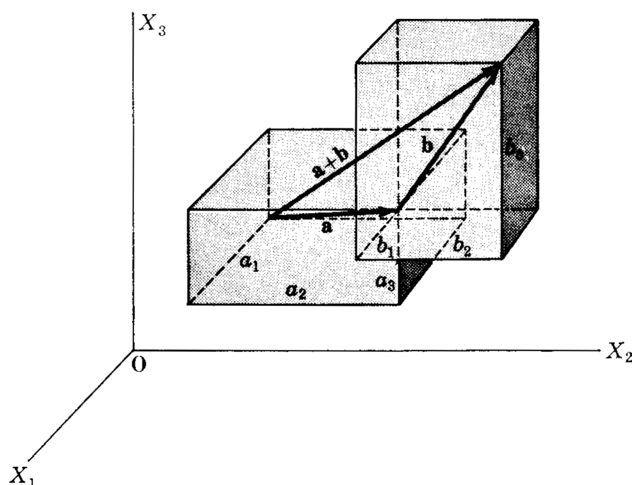


FIGURA 2

representen el mismo vector **a** serán de la misma longitud (magnitud) y apuntarán en la misma dirección. Es en este sentido que se dice que un vector especifica una "magnitud" y una "dirección".

La suma $\mathbf{a}+\mathbf{b} = (a_1+b_1, a_2+b_2, a_3+b_3)$ de un par de vectores en V_3 está ilustrada en la figura 2. El punto inicial de **b** se coloca en el punto terminal de **a**. La flecha **a** + **b** es entonces la flecha que tiene como punto inicial el de **a** y como punto terminal el de **b**.

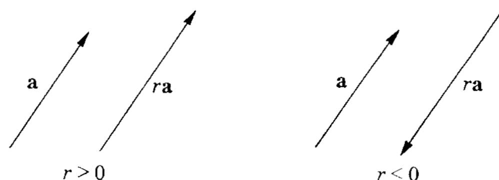


FIGURA 3

La figura 3 ilustra la multiplicación de un vector **a** por un número real r. La flecha $r\mathbf{a}$ es paralela a la flecha **a** y su longitud es $|r|$ veces la longitud de **a**; $r\mathbf{a}$ apunta en la misma dirección que **a** si $r > 0$, y si $r < 0$, la dirección de $r\mathbf{a}$ es la opuesta a la de **a**.

La figura 4 ilustra la ley conmutativa A_2 de la adición de vectores. La suma $\mathbf{a}+\mathbf{b}$ (figura 4) es una diagonal del paralelogramo cuyos lados son $\mathbf{a}$ y $\mathbf{b}$. La otra diagonal está relacionada con la diferencia de los dos

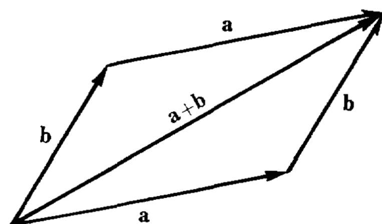


FIGURA 4

vectores. Esto se ilustra en la figura 5. Los vectores $\mathbf{a}$ y $\mathbf{b}$ están construidos con el mismo punto inicial. El vector $\mathbf{a}-\mathbf{b}$ es, entonces, el vector del punto terminal de $\mathbf{b}$ al punto terminal de $\mathbf{a}$. La figura 5 ilustra también que $\mathbf{b}+(\mathbf{a}-\mathbf{b}) = \mathbf{a}$. La ley asociativa A_3 se ilustra en la figura 6.

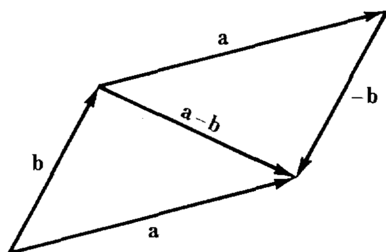


FIGURA 5

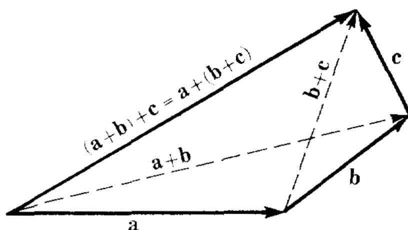


FIGURA 6

Problemas

1. Calcúlese gráficamente lo siguiente:

a) $(3, -5)+(5, -3)$ *b*) $(3, -5)-(2, 5)$
c) $(1, 1)+(-2, 5)+(-3, -2)$ *d*) $(1, 1)+(-2, 1)+(1, -2)$
e) $(\cos 30^\circ, \operatorname{sen} 30^\circ) + (\cos 45^\circ, \operatorname{sen} 45^\circ)$.

2. Demuéstrese gráficamente que hay números reales r y s que satisfacen

$$\mathbf{c} = r\mathbf{a}+s\mathbf{b}$$

donde

a) $\mathbf{a} = (5, 1)$, $\mathbf{b} = (3, 5)$, $\mathbf{c} = (5, 5)$
b) $\mathbf{a} = (2, -1)$, $\mathbf{b} = (3, 2)$, $\mathbf{c} = (5, 2)$
c) $\mathbf{a} = (-1, -2)$, $\mathbf{b} = (-1, 3)$, $\mathbf{c} = (4, 1)$

d) $\mathbf{a} = (-2, 3)$, $\mathbf{b} = (4, -1)$, $\mathbf{c} = (-3, 4)$
e) $\mathbf{a} = (1, 1, 1)$, $\mathbf{b} = (1, 0, 0)$, $\mathbf{c} = (4, 2, 2)$
f) $\mathbf{a} = (1, 1, 0)$, $\mathbf{b} = (-1, 2, 0)$, $\mathbf{c} = (3, 5, 0)$.

3. ¿Qué condiciones sobre **a**, **b** y **c** nos aseguran que **a**, **b** y **c** son los lados de un triángulo?

4. ¿Cuál es el significado geométrico de la ley distributiva S_5 del teorema 2.8?

4. PARALELISMO DE VECTORES

Al discutir la interpretación geométrica de la multiplicación de un vector por un número real, vimos que los vectores **a** y $r\mathbf{a}$, donde $r \neq 0$, están representados por flechas que son paralelas (figura 3). Definimos ahora el paralelismo entre vectores.

4.1 Definición. *Se dice que dos vectores en V_n son* ***paralelos*** *si uno de ellos es igual al producto del otro por un número real.*

Obsérvese que como $\mathbf{0} = 0\mathbf{a}$ para todo $\mathbf{a} \in V_n$, el vector cero es paralelo a todos los vectores.

4.2 Definición. *Dos vectores distintos de cero* **a** *y* **b** *en V_n se dice que tienen la* ***misma dirección*** *si* $\mathbf{b} = r\mathbf{a}$ *donde* $r > 0$, *y se dice que tienen* ***direcciones opuestas*** *si* $\mathbf{b} = r\mathbf{a}$ *donde* $r < 0$.

4.3 Ejemplo. ¿Son paralelos los vectores $(2, 1, 5)$ y $(-6, -3, -15)$?

SOLUCIÓN. Como $(-6, -3, -15) = -3(2, 1, 5)$, los vectores son paralelos y de direcciones opuestas.

4.4 Ejemplo. ¿Son paralelos los vectores $(1, 3, 2)$ y $(3, 9, 7)$?

SOLUCIÓN. Si los vectores fueran paralelos, como ninguno de ellos es cero, cada uno de ellos sería paralelo al otro y habría un número real r tal que

$$(1, 3, 2) = r(3, 9, 7).$$

Pero esto implica que $3r = 1$, $9r = 3$ y $7r = 2$, y no hay ningún número real r con esta propiedad. Por tanto, los vectores no son paralelos.

Problemas

1. ¿Cuáles de los siguientes pares de vectores están en la misma dirección?, ¿cuáles son paralelos?

a) (1, 1), (2, 2) *b*) (3, 8), (8, 24)
c) (1, 2, 1, −1), (−3, −6, −3, −3)
d) (1, −2, 2, −1), (−2, 4, −4, 2)
e) (5, 7, 2), (−15, −21, −6)
f) (3, 9), (−4, −6)

2. Pruébese que si $\mathbf{c} \neq \mathbf{0}$ y si **a** y **b** son paralelos a **c**, entonces **a** y **b** son paralelos. (Vectores paralelos a un mismo vector no nulo son paralelos entre sí.)

3. Pruébese que si $\mathbf{d} = \mathbf{b} + \mathbf{c}$ y si **b** es paralelo a **a**, entonces **d** es paralelo a **a** si y sólo si **c** es paralelo a **a**. Ilústrese este resultado gráficamente.

5. ORTOGONALIDAD DE VECTORES

Sea $\mathbf{a} = (a_1, a_2, a_3)$ un vector en V_3. En la sección 3 dimos una interpretación geométrica del vector **a** como si fuera una flecha en el espacio (figura 7). Si el espacio es euclidiano y si los ejes son rectangulares (mutuamente perpendiculares), entonces el teorema de Pitágoras se verifica y, por tanto, la longitud de la flecha que representa a **a** es $\sqrt{a_1^2 + a_2^2 + a_3^2}$. Como nuestra geometría tiene que ser euclidiana, definimos la longitud del vector **a** en V_3 como $\sqrt{a_1^2 + a_2^2 + a_3^2}$. Generalizando esta longitud euclidiana, introducimos la siguiente

5.1 Definición. *La longitud de un vector* $\mathbf{a} = (a_1, \ldots, a_n) \in V_n$, *denotada por* $|\mathbf{a}|$, *tiene como definición*

$$|\mathbf{a}| = \sqrt{a_1^{\,2} + \ldots + a_n^{\,2}} = \left[\sum_{k=1}^{n} a_k^{\,2} \right]^{1/2}$$

Un vector de longitud igual a la unidad se llama vector unitario. A V_n con la longitud que acabamos de definir se le llama espacio vectorial euclidiano n-dimensional.

5.2 Teorema. ***Las propiedades fundamentales de la longitud de un vector son:*** *Para* $\mathbf{a}$, $\mathbf{b} \in V_n$ *cualesquiera y para todo* $r \in \mathrm{R}$,

5.3 $|\mathbf{a}| \geqslant 0$; $|\mathbf{a}| = 0$ *si y sólo si* $\mathbf{a} = \mathbf{0}$.

5.4 $|r\mathbf{a}| = |r|\,|\mathbf{a}|$.

5.5 $|\mathbf{a} + \mathbf{b}| \leqslant |\mathbf{a}| + |\mathbf{b}|$ (desigualdad del triángulo).

PRUEBA DE 5.3. Por definición, $|\mathbf{a}| \geqslant 0$. Ahora bien $|\mathbf{a}|^2 = a_1^{\,2} + \ldots + a_n^{\,2}$, y por tanto, si $a_i \neq 0$ para un cualquier $i = 1, \ldots, n$, entonces $|\mathbf{a}| \neq 0$. Por tanto $|\mathbf{a}| = 0$ implica $a_1 = 0, \ldots, a_n = 0$, luego

$$\mathbf{a} = (a_1, \ldots, a_n) = (0, \ldots, 0) = \mathbf{0}.$$

Recíprocamente, si $\mathbf{a} = \mathbf{0}$, entonces $|\mathbf{a}| = 0$.

PRUEBA DE 5.4

$$|r\mathbf{a}| = |(ra_1, \ldots, ra_n)| = \sqrt{(ra_1)^2 + \ldots + (ra_n)^2}$$
$$= \sqrt{r^2} \cdot \sqrt{a_1^2 + \ldots + a_n^2}$$
$$= \sqrt{r^2} \cdot |\mathbf{a}| = |r|\,|\mathbf{a}|.$$

La prueba de la desigualdad del triángulo se da en la sección 7 (pág. 35). Entonces no nos será difícil demostrar que, si $\mathbf{a}$ y $\mathbf{b}$ no están en la misma dirección, entonces $|\mathbf{a} + \mathbf{b}| < |\mathbf{a}| + |\mathbf{b}|$, $(\mathbf{a} \neq \mathbf{0}, \mathbf{b} \neq \mathbf{0})$. Esta desigualdad corresponde al teorema geométrico: *la longitud de un lado de un triángulo no degenerado es menor que la suma de las longitudes de los otros dos lados* (figura 8).

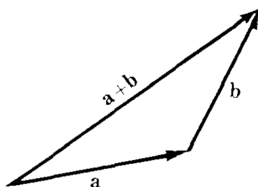


FIGURA 8

Adviértase que la notación para la longitud de un vector es la misma que la usada para el valor absoluto de un número real. La razón para haber elegido tal notación es que las propiedades fundamentales del valor absoluto de un número real y las de la longitud de un vector son las mismas. En realidad, si consideramos a los números reales como vectores en V_1, entonces el valor absoluto es la longitud del vector unidimensional, es decir, $|r| = \sqrt{r^2}$.

Volviendo a nuestra imagen geométrica de los vectores, queremos motivar la definición que acabamos de dar. La palabra "ortogonal" significa "en ángulo recto" y es sinónima de "perpendicular". Sean **a** y **b** los lados de un paralelogramo (figura 9). Los vectores **a** + **b** y **a** − **b** son las

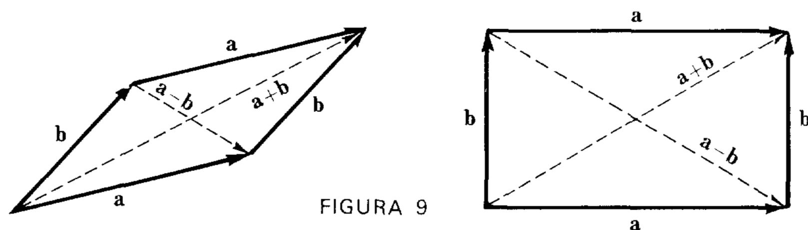


FIGURA 9

diagonales del paralelogramo. Expresada geométricamente, la definición de ortogonalidad podría ser: **a** es "ortogonal" a **b** si las diagonales del paralelogramo formado por **a** y **b** son de igual longitud, es decir, si el paralelogramo es un rectángulo.

5.6 Definición. *Un vector* **a** *se dice que es* ***ortogonal*** *a un vector* **b** *si*

$$|\mathbf{a}+\mathbf{b}| = |\mathbf{a}-\mathbf{b}|.$$

Como $|\mathbf{a}+\mathbf{b}| = |\mathbf{b}+\mathbf{a}|$ y $|\mathbf{a}-\mathbf{b}| = |\mathbf{b}-\mathbf{a}|$, es claro que **a** ortogonal a **b** implica **b** ortogonal a **a**. Por esta razón se usa con frecuencia la expresión "mutuamente ortogonales". Diremos también que, a veces, "**a** y **b** son ortogonales".

El vector cero tiene la propiedad muy especial de ser ortogonal a todos los vectores.

5.7 Ejemplo. ¿Son ortogonales los vectores $\mathbf{a} = (5, -8, 3)$ y $\mathbf{b} = (2, 5, 10)$?

SOLUCIÓN

$$\begin{aligned} |\mathbf{a}+\mathbf{b}| &= |(5, -8, 3)+(2, 5, 10)| = |(7, -3, 13)| \\ &= \sqrt{7^2+(-3)^2+13^2} = \sqrt{227} \\ |\mathbf{a}-\mathbf{b}| &= |(5, -8, 3)-(2, 5, 10)| = |(3, -13, -7)| \\ &= \sqrt{3^2+(-13)^2+(-7)^2} = \sqrt{227}. \end{aligned}$$

Como $|\mathbf{a}+\mathbf{b}| = |\mathbf{a}-\mathbf{b}|$, los vectores son ortogonales.

5.8 Ejemplo. ¿Los vectores $\mathbf{a} = (-2, 6, 4, -3)$ y $\mathbf{b} = (3, \frac{1}{2}, 1, -1)$ son ortogonales?

SOLUCIÓN

$$|\mathbf{a}+\mathbf{b}| = |(1, \tfrac{13}{2}, 5, -4)| = \sqrt{1+\tfrac{169}{4}+25+16} = \sqrt{42+\tfrac{169}{4}}$$

$$|\mathbf{a}-\mathbf{b}| = |(-5, \tfrac{11}{2}, 3, -2)| = \sqrt{25+\tfrac{121}{4}+9+4} = \sqrt{38+\tfrac{121}{4}}.$$

Como $|\mathbf{a}+\mathbf{b}| \neq |\mathbf{a}-\mathbf{b}|$, los vectores no son ortogonales.

Problemas

1. Si $\mathbf{a} = (3, 0, 5)$, $\mathbf{b} = (2, -1, -3)$, calcúlese la longitud de:

a) $\mathbf{a}$ *b*) $\mathbf{b}$ *c*) $\mathbf{a}+\mathbf{b}$ *d*) $\mathbf{a}-\mathbf{b}$
e) $3\mathbf{a}$ *f*) $-3(\mathbf{a}-\mathbf{b})$ *g*) $-\frac{1}{2}\mathbf{b}$ *h*) $3\mathbf{a}-\frac{1}{2}\mathbf{b}$
i) $\mathbf{a}/|\mathbf{a}|$ *j*) $\mathbf{b}/|\mathbf{b}|$ *k*) $(\mathbf{a}+\mathbf{b})/|\mathbf{a}+\mathbf{b}|$ *l*) $\dfrac{\mathbf{a}}{|\mathbf{a}|} + \dfrac{\mathbf{b}}{|\mathbf{b}|}$

2. Demuéstrese que $|-\mathbf{a}| = |\mathbf{a}|$.

3. Demuéstrese que $\big||\mathbf{a}|-|\mathbf{b}|\big| \leqslant |\mathbf{a}-\mathbf{b}|$ para todo $\mathbf{a}, \mathbf{b} \in V_n$.

4. Determínese si los siguientes pares de vectores son ortogonales.

a) $(-1, 3, -3)$ y $(3, 3, 2)$
b) $(1, 0, 0)$ y $(0, 1, 0)$
c) $(2, 8, 4)$ y $(0, 0, 0)$
d) $(3, 2, 0)$ y $(1, -1, 0)$

5. *a*) Demuéstrese que $\mathbf{a}+\mathbf{b}$ y $\mathbf{a}-\mathbf{b}$ son ortogonales si y sólo si $|\mathbf{a}| = |\mathbf{b}|$.
b) ¿Cuál es la interpretación geométrica del problema 5*a*?

6. ¿Qué es lo que puede concluirse si se sabe que un vector es perpendicular a sí mismo?

7. Encuéntrense todos los vectores ortogonales a

a) $(3, 6)$ *b*) $(2, -1)$
c) (a_1, a_2) *d*) $(2, 3, -1)$

8. Demuéstrese que si $\mathbf{a}$ es un vector distinto de cero, entonces $\dfrac{1}{|\mathbf{a}|}\mathbf{a}$ es un vector de longitud igual a uno que está en la dirección de $\mathbf{a}$. A un vector de longitud igual a la unidad se llama *vector unitario*.

9. Encuéntrense los vectores unitarios en la dirección de:

a) $(1, 1)$ *b*) $(1, -1, 1)$ *c*) $(2, 3, -7)$

10. Demuéstrese que si **a** y **b** son vectores distintos de cero, entonces **a** ortogonal a **b** implica que **a** no es paralelo a **b**, y, recíprocamente, **a** paralelo a **b** implica que **a** no es ortogonal a **b**.

6. EL PRODUCTO ESCALAR

Nuestra definición de ortogonalidad de un par de vectores $\mathbf{a} = (a_1, \ldots, a_n)$ y $\mathbf{b} = (b_1, \ldots, b_n)$ es equivalente a afirmar que la diferencia de los cuadrados de las longitudes de las diagonales $\mathbf{a}+\mathbf{b}$ y $\mathbf{a}-\mathbf{b}$ del paralelogramo de lados **a** y **b** es cero; es decir,

$$|\mathbf{a}+\mathbf{b}|^2 - |\mathbf{a}-\mathbf{b}|^2 = 0.$$

Como

6.1 $$|\mathbf{a}+\mathbf{b}|^2 - |\mathbf{a}-\mathbf{b}|^2 = \sum_{k=1}^{n} (a_k + b_k)^2 - \sum_{k=1}^{n} (a_k - b_k)^2$$
$$= \sum_{k=1}^{n} (a_k{}^2 + 2a_k b_k + b_k{}^2 - a_k{}^2 + 2a_k b_k - b_k{}^2)$$
$$= 4 \sum_{k=1}^{n} a_k b_k,$$

la ortogonalidad de los dos vectores **a** y **b** es equivalente a la anulación de $\sum_{k=1}^{n} a_k b_k$. Esta expresión $\sum_{k=1}^{n} a_k b_k$ es de considerable importancia en álgebra, geometría y física, y es por ello que se le ha dado un nombre especial.

6.2 Definición. *El producto escalar* $\mathbf{a}\cdot\mathbf{b}$ *—léase "*__a__ *producto escalar* __b__*", "*__a__ *punto* __b__*", o simplemente, "*__a__ *por* __b__*"— de dos vectores* __a__, $\mathbf{b} \in V_n$, *donde* $\mathbf{a} = (a_1, \ldots, a_n)$ *y* $\mathbf{b} = (b_1, \ldots, b_n)$, *está definido por*

$$\mathbf{a}\cdot\mathbf{b} = \sum_{k=1}^{n} a_k b_k = a_1 b_1 + \ldots + a_n b_n .$$

Nótese que el producto escalar de dos vectores *no es* un vector, *es un número real.* En física, magnitudes tales como la longitud, el trabajo, la masa, la temperatura, etc., se llaman magnitudes "escalares"; tienen magnitud, pero no dirección y quedan especificadas (medidas) por números reales. En matemáticas, a menudo se usa el término "producto interior" en lugar del término "producto escalar". Otro nombre para este producto —sugerido por la notación— es el de "producto punto".

La ecuación 6.1 puede escribirse ahora:

6.3 $$|\mathbf{a}+\mathbf{b}|^2 - |\mathbf{a}-\mathbf{b}|^2 = 4(\mathbf{a}\cdot\mathbf{b});$$

y podemos enunciar:

6.4 Teorema. *Dos vectores* **a** *y* **b** *son ortogonales si y sólo si* $\mathbf{a} \cdot \mathbf{b} = 0$.

6.5 Ejemplo. Aplíquese el criterio que acaba de enunciarse para estudiar la ortogonalidad en los ejemplos 5.7 y 5.8.

SOLUCIÓN DE 5.7

$$\mathbf{a} \cdot \mathbf{b} = (5, -8, 3) \cdot (2, 5, 10) = 10 - 40 + 30 = 0.$$

Por tanto, los vectores son ortogonales.

SOLUCIÓN DE 5.8

$$\mathbf{a} \cdot \mathbf{b} = (-2, 6, 4, -3) \cdot (3, \tfrac{1}{2}, 1, -1) = -6 + 3 + 4 + 3 = 4 \neq 0.$$

Por tanto, los vectores no son ortogonales.

6.6 Teorema. ***Las propiedades fundamentales del producto escalar son:***

6.7 $$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$$

6.8 $$(r\mathbf{a}) \cdot \mathbf{b} = r(\mathbf{a} \cdot \mathbf{b})$$

6.9 $$\mathbf{a} \cdot (\mathbf{b}_1 + \mathbf{b}_2) = \mathbf{a} \cdot \mathbf{b}_1 + \mathbf{a} \cdot \mathbf{b}_2$$

6.10 $$\mathbf{a} \cdot \mathbf{a} \geqslant 0; \quad \mathbf{a} \cdot \mathbf{a} = 0 \quad \textit{si y sólo si} \quad \mathbf{a} = \mathbf{0}.$$

La propiedad 6.7 afirma que la ley conmutativa se verifica para el producto escalar y la 6.9 que también se cumple la ley distributiva. La propiedad 6.10 se ve que es una reformulación de la propiedad 5.3 de $|\mathbf{a}|$ ya que

6.11 $$\mathbf{a} \cdot \mathbf{a} = \sum_{k=1}^{n} a_k{}^2 = |\mathbf{a}|^2.$$

La propiedades 6.7, 6.8 y 6.9, son simples consecuencias de las propiedades de los números reales (problema 4).

Mostramos ahora que las definiciones de longitud y ortogonalidad implican el teorema de Pitágoras.

6.12 Teorema. **a** *es ortogonal a* **b** *si y sólo si*

$$|\mathbf{a} + \mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2.$$

PRUEBA. De acuerdo con las propiedades fundamentales del producto escalar,

$$\begin{aligned} |\mathbf{a} + \mathbf{b}|^2 &= (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) \\ &= \mathbf{a} \cdot (\mathbf{a} + \mathbf{b}) + \mathbf{b} \cdot (\mathbf{a} + \mathbf{b}) \\ &= \mathbf{a} \cdot \mathbf{a} + \mathbf{a} \cdot \mathbf{b} + \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} \\ &= |\mathbf{a}|^2 + 2\,\mathbf{a} \cdot \mathbf{b} + |\mathbf{b}|^2. \end{aligned}$$

Vemos pues que $|\mathbf{a}+\mathbf{b}|^2 = |\mathbf{a}|^2+|\mathbf{b}|^2$ si y sólo si $\mathbf{a}\cdot\mathbf{b} = 0$; es decir, si y sólo si **a** es ortogonal a **b**.

Problemas

1. Sean $\mathbf{a} = (3, 0, 5)$, $\mathbf{b} = (2, -1, -3)$, $\mathbf{P}_0 = (1, -2, 1)$, y $\mathbf{P}_1 = (2, 3, -1)$. Encuéntrense

a) $\mathbf{a}\cdot\mathbf{b}$
b) $\mathbf{a}\cdot(\mathbf{P}_1-\mathbf{P}_0)$
c) $\mathbf{b}\cdot(\mathbf{P}_1-\mathbf{P}_0)$
d) $(\mathbf{a}+\mathbf{b})\cdot(\mathbf{P}_1-\mathbf{P}_0)$
e) $\mathbf{a}\cdot\mathbf{a}$
f) $\mathbf{b}\cdot\mathbf{b}$
g) $(\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}-\mathbf{b})$
h) $|\mathbf{a}+\mathbf{b}|^2$

2. Determínese si los siguientes pares de vectores son ortogonales.

a) $(2, 1, -3, 4)$ y $(3, 4, 2, -1)$
b) $(3, 2, 0, -1)$ y $(4, -1, 7, 2)$
c) $(-18, 2, 3, 4)$ y $(2, 6, 12, 3)$
d) $(1, 0, 0, 0)$ y $(0, 0, 1, 0)$

3. Encuéntrense todos los vectores ortogonales a:

a) $(3, 6)$
b) $(2, -1)$
c) $(1, 0, 0)$ y $(0, 1, 0)$
d) $(1, 1, 1)$ y $(0, 0, 1)$
e) (a_1, a_2)
f) (a_1, a_2, a_3)

4. Pruébese que el producto escalar satisface 6.7, 6.8 y 6.9.

5. Demuéstrese que:

a) $|\mathbf{a}+\mathbf{b}|^2 = |\mathbf{a}|^2+2\mathbf{a}\cdot\mathbf{b}+|\mathbf{b}|^2$
b) $(\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}-\mathbf{b}) = |\mathbf{a}|^2-|\mathbf{b}|^2$
c) $|\mathbf{a}+t\mathbf{b}|^2 = |\mathbf{a}|^2+2t\mathbf{a}\cdot\mathbf{b}+t^2|\mathbf{b}|^2$

6. Pruébese que la suma de los cuadrados de las longitudes de las diagonales de un paralelogramo es igual a la suma de los cuadrados de las longitudes de los cuatro lados del paralelogramo.

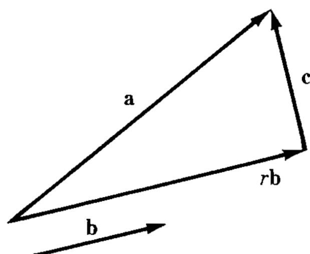
7. PROYECCIÓN ORTOGONAL. COMPONENTES

En esta sección discutiremos la significación geométrica del producto escalar en términos de "proyección ortogonal" y "componente". Estos conceptos son de importancia tanto en geometría como en física.

Introducimos los conceptos de proyección ortogonal y componente en conexión con el siguiente problema. Dados dos vectores no nulos **a** y **b** constrúyase un triángulo rectángulo con hipotenusa **a** y base paralela a **b** (figura 10). Como cualquier vector paralelo a **b** puede representarse por

$r\mathbf{b}$ con r igual a algún número real, lo que deseamos es construir un triángulo de lados $\mathbf{a}$, $r\mathbf{b}$ y $\mathbf{c} = \mathbf{a} - r\mathbf{b}$ tal que $\mathbf{c}$ sea ortogonal a $\mathbf{b}$. Pero $\mathbf{a} - r\mathbf{b}$ es ortogonal a $\mathbf{b}$ si y sólo si

$$(\mathbf{a} - r\mathbf{b}) \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{b} - r|\mathbf{b}|^2 = 0.$$



$\mathbf{a} = r\mathbf{b} + \mathbf{c}$

FIGURA 10

Por tanto, $r = \dfrac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|^2}$ es el único número tal que $\mathbf{a} - r\mathbf{b}$ es ortogonal a $\mathbf{b}$ y el triángulo rectángulo deseado de hipotenusa $\mathbf{a}$ tiene lados $\dfrac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|^2}\mathbf{b}$ y $\mathbf{a} - \dfrac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|^2}\mathbf{b}$. El lado $\dfrac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|^2}\mathbf{b}$ que es paralelo a $\mathbf{b}$ se llama proyección ortogonal de $\mathbf{a}$ sobre $\mathbf{b}$.

7.1 Definición. *Sean* $\mathbf{a}, \mathbf{b} \in V_n$ *con* $\mathbf{b} \neq 0$. *La* ***proyección ortogonal*** *de* $\mathbf{a}$ *sobre* $\mathbf{b}$, *denotada por* $\mathbf{Proy_b\, a}$, *es el vector*

$$\mathbf{Proy_b\, a} = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|^2}\mathbf{b}.$$

La proyección de $\mathbf{a}$ sobre $\mathbf{b}$ puede escribirse en la forma

$$\mathbf{Proy_b\, a} = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|}\frac{\mathbf{b}}{|\mathbf{b}|}.$$

Como el vector $\dfrac{\mathbf{b}}{|\mathbf{b}|}$ es un vector unitario en la dirección de $\mathbf{b}$, el número $\dfrac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|}$ es la "longitud dirigida" de $\mathbf{Proy_b\, a}$. Este número se llama componente de $\mathbf{a}$ en la dirección de $\mathbf{b}$.

7.2 Definición. *El número* $(\mathbf{a} \cdot \mathbf{b})/|\mathbf{b}|$ *se llama* ***componente*** *de* $\mathbf{a}$ *en la dirección de* $\mathbf{b}$ *y se denota por* $\mathrm{Comp}_\mathbf{b}\, \mathbf{a}$; *es decir*,

$$\mathrm{Comp}_\mathbf{b}\, \mathbf{a} = (\mathbf{a} \cdot \mathbf{b})/|\mathbf{b}|.$$

La relación entre proyección (un vector) y componente (un número) es

7.3 $$\mathbf{Proy_b\ a} = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{b}|^2}\mathbf{b} = (\mathrm{Comp_b}\ \mathbf{a})\frac{\mathbf{b}}{|\mathbf{b}|}.$$

Si $\mathrm{Comp_b}\ \mathbf{a} > 0$, entonces $\mathbf{Proy_b\ a}$ está en la dirección de $\mathbf{b}$ (figura 11a). Si $\mathrm{Comp_b}\ \mathbf{a} < 0$, entonces $\mathbf{Proy_b\ a}$ y $\mathbf{b}$ están en direcciones opuestas (figura 11b). Si $\mathrm{Comp_b}\ \mathbf{a} = 0$, entonces los vectores $\mathbf{a}$ y $\mathbf{b}$ son ortogonales.

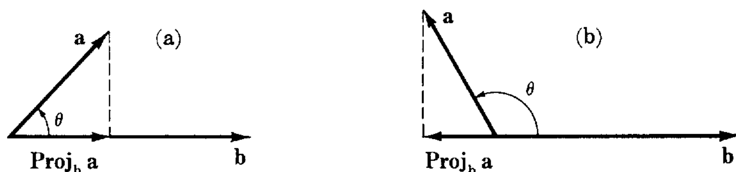


FIGURA 11

Nota. No hay mucha concordancia entre los distintos autores respecto a la terminología de componentes y proyecciones. Algunos autores usan el término "componente" tanto para el vector al que nosotros hemos designado como $\mathbf{Proy_b\ a}$ como para el número al que hemos denotado como $\mathrm{Comp_b}\ \mathbf{a}$. Cuando se hace esto, es común hablar de "componente vectorial" y de "componente escalar" cuando se necesita distinguir entre los dos conceptos. Otros autores usan tanto el término "componente" como el término "proyección" para denotar el número que aquí hemos denominado $\mathrm{Comp_b}\ \mathbf{a}$.

Nota. Si $\mathbf{b}'$ es un vector cualquiera no nulo paralelo a $\mathbf{b}$, entonces $\mathbf{Proy_b\ a} = \mathbf{Proy_{b'}\ a}$ (problema 5*a*). Así pues $\mathbf{Proy_b\ a}$ no cambia porque reemplacemos $\mathbf{b}$ por cualquier vector no nulo paralelo a $\mathbf{b}$. Por otra parte, si $\mathbf{b}'$ es un vector distinto de cero paralelo a $\mathbf{b}$, entonces $\mathrm{Comp_{b'}}\ \mathbf{a} = \mathrm{Comp_b}\ \mathbf{a}$ o $\mathrm{Comp_{b'}}\ \mathbf{a} = -\mathrm{Comp_b}\ \mathbf{a}$ según que $\mathbf{b}$ y $\mathbf{b}'$ tengan igual dirección o direcciones opuestas (problemas 5*b* y 5*c*).

Como el componente de un vector en la dirección de otro vector tiene un significado geométrico definido, la relación entre componente y producto escalar introduce una interpretación geométrica del producto escalar. Según la definición de componente (definición 7.2),

7.4 $$\mathbf{a}\cdot\mathbf{b} = |\mathbf{b}|\ \mathrm{Comp_b}\ \mathbf{a}.$$

Esta ecuación nos dice: *el producto escalar* $\mathbf{a}\cdot\mathbf{b}$ *es la longitud de* $\mathbf{b}$ *por el componente de* $\mathbf{a}$ *en la dirección de* $\mathbf{b}$.

En el espacio vectorial bidimensional V_2 (figura 11),

$$\mathrm{Comp_b}\ \mathbf{a} = |\mathbf{a}| \cos \theta,$$

donde θ es el ángulo de **b** a **a** y, por tanto,

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos \theta .$$

La misma terminología puede extenderse para V_n. Consideramos un "ángulo" en el espacio n-dimensional como determinado por un par de vectores distintos de cero **a** y **b**. Si θ es el ángulo determinado por los vectores no nulos **a** y **b** en V_n, definimos el coseno de θ por la relación

7.5
$$\cos \theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} .$$

La interpretación geométrica del producto escalar sugiere una importante propiedad llamada desigualdad de Schwarz.

7.6 Teorema. (Desigualdad de Schwarz.) *Para cualquier* **a**, **b** $\in V_n$,

7.7
$$|\mathbf{a} \cdot \mathbf{b}| \leqslant |\mathbf{a}|\,|\mathbf{b}|$$

donde la igualdad se verifica si y sólo si **a** *y* **b** *son paralelos.*

PRUEBA. (Figura 10, pág. 32.) Si **a** no es paralela a **b**, hay un triángulo rectángulo con hipotenusa **a** y base $\mathbf{Proy_b\, a}$. Sea $\mathbf{c} = \mathbf{a} - \mathbf{Proy_b\, a} \neq \mathbf{0}$ el tercer lado de este triángulo rectángulo. De acuerdo con el teorema de Pitágoras tenemos

$$|\mathbf{Proy_b\, a}|^2 = |\mathbf{a}|^2 - |\mathbf{c}|^2 < |\mathbf{a}|^2$$

o

$$|\mathbf{Proy_b\, a}| < |\mathbf{a}| .$$

De las ecuaciones 7.4 y 7.3 se deduce

$$|\mathbf{a} \cdot \mathbf{b}| = |\mathbf{b}|\,|\mathrm{Comp}_\mathbf{b}\, \mathbf{a}| = |\mathbf{b}|\,|\mathbf{Proy_b\, a}| < |\mathbf{b}|\,|\mathbf{a}|$$

de modo que si **a** no es paralelo a **b**,

$$|\mathbf{a} \cdot \mathbf{b}| < |\mathbf{a}|\,|\mathbf{b}| .$$

Si **a** es paralelo a **b** y **a** o **b** es nulo, entonces la igualdad se verifica $(0 = 0)$. Si **a** y **b** son vectores paralelos no nulos, entonces $\mathbf{a} = r\mathbf{b}$ para algún número real r y, por tanto,

$$|\mathbf{a} \cdot \mathbf{b}| = |(r\mathbf{b}) \cdot \mathbf{b}| = |r|\,|\mathbf{b}|^2 = |r\mathbf{b}|\,|\mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}| .$$

Esto completa la prueba.

Nota. Como una consecuencia inmediata de la desigualdad de Schwarz se sigue que $\cos \theta$, de acuerdo a como ha sido definida por la ecuación 7.5, satisface la desigualdad $-1 \leqslant \cos \theta \leqslant 1$.

La desigualdad del triángulo 5.5 es ahora una simple consecuencia de la desigualdad de Schwarz.

PRUEBA DE LA DESIGUALDAD DEL TRIÁNGULO 5.5. Como $\mathbf{a} \cdot \mathbf{b} \leqslant |\mathbf{a} \cdot \mathbf{b}|$, de la desigualdad de Schwarz se deduce

$$\mathbf{a} \cdot \mathbf{b} \leqslant |\mathbf{a} \cdot \mathbf{b}| \leqslant |\mathbf{a}|\,|\mathbf{b}|.$$

De donde

$$\begin{aligned} |\mathbf{a}+\mathbf{b}|^2 &= (\mathbf{a}+\mathbf{b}) \cdot (\mathbf{a}+\mathbf{b}) = |\mathbf{a}|^2 + 2\mathbf{a} \cdot \mathbf{b} + |\mathbf{b}|^2 \\ &\leqslant |\mathbf{a}|^2 + 2|\mathbf{a}|\,|\mathbf{b}| + |\mathbf{b}|^2 = (|\mathbf{a}| + |\mathbf{b}|)^2. \end{aligned}$$

Esto implica la desigualdad del triángulo:

$$|\mathbf{a}+\mathbf{b}| \leqslant |\mathbf{a}| + |\mathbf{b}|.$$

Es claro que si **a** es cero o lo es **b**, entonces se verifica la igualdad en la desigualdad del triángulo. Si **a** y **b** son distintos de cero, la igualdad se verifica si y solamente si

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a} \cdot \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|.$$

La igualdad se verifica en la desigualdad de Schwarz ($|\mathbf{a} \cdot \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|$), si y sólo si **a** y **b** son paralelas, es decir, $\mathbf{a} = r\mathbf{b}$ para un cierto número real r. Si $\mathbf{a} = r\mathbf{b}$, entonces

$$|\mathbf{a}|\,|\mathbf{b}| = |r\mathbf{b}|\,|\mathbf{b}| = |r|\,|\mathbf{b}|^2$$

y

$$\mathbf{a} \cdot \mathbf{b} = (r\mathbf{b}) \cdot \mathbf{b} = r|\mathbf{b}|^2$$

y por tanto $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}|$ si y sólo si $r = |r|$, es decir, $r \geqslant 0$. Vemos pues que la igualdad se verifica en la desigualdad del triángulo si y sólo si $\mathbf{a} = \mathbf{0}$, $\mathbf{b} = \mathbf{0}$, o **a** y **b** están en la misma dirección.

Problemas

1. Exprésese, en cada uno de los siguientes casos, **a** como la suma de un vector paralelo a **b** y un vector ortogonal a **b**.

a) $\mathbf{a} = (3, 8)$, $\mathbf{b} = (1, 0)$ *b*) $\mathbf{a} = (1, 0)$, $\mathbf{b} = (3, 8)$
c) $\mathbf{a} = (-5, 8)$, $\mathbf{b} = (1, 1)$ *d*) $\mathbf{a} = (1, 2, 3)$, $\mathbf{b} = (0, 0, 1)$
e) $\mathbf{a} = (1, 2, 3)$, $\mathbf{b} = (1, 1, 0)$ *f*) $\mathbf{a} = (2, 1, 1)$, $\mathbf{b} = (1, 2, 0)$

2. Ilústrense gráficamente las soluciones del problema 1.

3. En cada uno de los siguientes casos calcúlense $\text{Comp}_{\mathbf{b}}\, \mathbf{a}$ y $\mathbf{Proy}_{\mathbf{b}}\, \mathbf{a}$.

a) $\mathbf{a} = (3, 8)$, $\mathbf{b} = (1, 0)$ *b*) $\mathbf{a} = (-5, 8)$, $\mathbf{b} = (1, 1)$
c) $\mathbf{a} = (1, 2, -3)$, $\mathbf{b} = (0, 0, 1)$ *d*) $\mathbf{a} = (1, 1, 1)$, $\mathbf{b} = (1, 0, 1)$

e) $\mathbf{a} = (1, 0, 1)$, $\mathbf{b} = (1, 1, 1)$ *f*) $\mathbf{a} = (1, 2, -3, 6)$, $\mathbf{b} = (1, 0, 1, 0)$
g) $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (0, a_2, 0)$

4. Demuéstrese que $\text{Comp}_{\mathbf{b}}(\mathbf{a}_1 + \mathbf{a}_2) = \text{Comp}_{\mathbf{b}}\,\mathbf{a}_1 + \text{Comp}_{\mathbf{b}}\,\mathbf{a}_2$ (la componente de una suma es la suma de las componentes). Ilústrese este resultado gráficamente.

5. Demuéstrese que:

a) Si $\mathbf{b}$ y $\mathbf{b}'$ son vectores paralelos no nulos, entonces $\mathbf{Proy}_{\mathbf{b}}\,\mathbf{a} = \mathbf{Proy}_{\mathbf{b}'}\,\mathbf{a}$.
b) Si $\mathbf{b}$ y $\mathbf{b}'$ están en la misma dirección, entonces $\text{Comp}_{\mathbf{b}}\,\mathbf{a} = \text{Comp}_{\mathbf{b}'}\,\mathbf{a}$.
c) Si $\mathbf{b}$ y $\mathbf{b}'$ están en direcciones opuestas, entonces $\text{Comp}_{\mathbf{b}}\,\mathbf{a} = -\text{Comp}_{\mathbf{b}'}\,\mathbf{a}$.

6. Obténgase una nueva demostración de la desigualdad de Schwarz mediante la consideración de la expresión $\left|\dfrac{\mathbf{a}}{|\mathbf{a}|} \pm \dfrac{\mathbf{b}}{|\mathbf{b}|}\right|^2$ para vectores no nulos $\mathbf{a}$ y $\mathbf{b}$.

8. VECTORES SOBRE UN CAMPO ARBITRARIO

En la sección 2, el espacio vectorial n-dimensional V_n fue definido como el conjunto de todas las n-adas de números reales con la relación de igualdad y las operaciones de adición y multiplicación por un escalar (número real) definidas como sigue:

8.1 Igualdad de vectores. Si $\mathbf{x} = (x_1, \ldots, x_n)$ y $\mathbf{y} = (y_1, \ldots, y_n)$ son vectores, entonces

$$\mathbf{x} = \mathbf{y} \quad \text{si} \quad x_i = y_i \quad \text{para todo} \quad i = 1, \ldots, n.$$

8.2 Adición de vectores. Si $\mathbf{x} = (x_1, \ldots, x_n)$ y $\mathbf{y} = (y_1, \ldots, y_n)$ son vectores, entonces

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n).$$

8.3 Multiplicación de un vector por un escalar. Si $\mathbf{x} = (x_1, \ldots, x_n)$ es un vector y r es un escalar, entonces

$$r\mathbf{x} = (rx_1, \ldots, rx_n).$$

Observemos ahora que podemos definir una estructura a la que llamaremos $V_n(F)$ si, en las anteriores definiciones, reemplazamos los números reales por elementos de algún conjunto F con tal de que los elementos de F puedan sumarse y multiplicarse. Sin embargo, para que podamos llamar a $V_n(F)$ espacio vectorial, exigimos que $V_n(F)$ tengan las propiedades que aparecen enumeradas en el teorema 2.8, pág. 3. Para

asegurarnos de que $V_n(F)$ tenga tales propiedades, suponemos que F es un campo.

Un *campo* es un conjunto F y dos operaciones, adición y multiplicación, que satisfacen las siguientes propiedades:

A_1. Para todo a y b en F, $a+b \in F$.
A_2. Para todo a y b en F, $a+b = b+a$.
A_3. Para todo a, b y c en F, $(a+b)+c = a+(b+c)$.
A_4. Hay un elemento en F, denotado por 0, tal que para todo a en F, $a+0 = a$.
A_5. Para cada a en F, hay un elemento en F, representado por $-a$, tal que $a+(-a) = 0$.
M_1. Para todo a y b en F, $ab \in F$.
M_2. Para todo a y b en F, $ab = ba$.
M_3. Para todo a, b y c en F, $(ab)c = a(bc)$.
M_4. Hay un elemento en F, representado por 1, diferente de 0, tal que para todo a en F, $a \cdot 1 = a$.
M_5. Para cada a en F, distinto de 0, hay un elemento en F, representado poi a^{-1}, tal que $a \cdot a^{-1} = 1$.
D. Para todo a, b y c en F, $a(b+c) = ab+ac$.

Definimos ahora el espacio vectorial n-dimensional $V_n(F)$ como el conjunto de todas las n-adas de elementos del campo F, denotadas por $\mathbf{x} = (x_1, \ldots, x_n)$, $x_i \in F$ $(i = 1, \ldots, n)$ y llamadas vectores, donde la relación de igualdad y las operaciones de adición y multiplicación por un escalar (elemento de F) satisfacen 8.1, 8.2 y 8.3. Como las únicas propiedades de los números reales que intervienen en la prueba del teorema 2.8 son las propiedades de campo, $V_n(F)$ tendrá también estas propiedades fundamentales.

8.4 Teorema.

A_1. *Para todo* $\mathbf{x}$ y $\mathbf{y}$ *en* $V_n(F)$, $\mathbf{x}+\mathbf{y} \in V_n(F)$.
A_2. *Para todo* $\mathbf{x}$ *y* $\mathbf{y}$ *en* $V_n(F)$, $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$.
A_3. *Para todo* $\mathbf{x}$, $\mathbf{y}$ *y* $\mathbf{z}$ *en* $V_n(F)$, $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$.
A_4. *Hay un y solamente un vector en* $V_n(F)$, *denotado por* $\mathbf{0}$ *y llamado vector cero, con la propiedad de que*

$$\mathbf{x}+\mathbf{0} = \mathbf{x} \quad \textit{para todo} \quad \mathbf{x} \in V_n(F).$$

A_5. *Para cada* $\mathbf{x} \in V_n(F)$ *hay un vector único, denotado por* $-\mathbf{x}$, *con la propiedad de que*

$$\mathbf{x}+(-\mathbf{x}) = \mathbf{0}.$$

S_1. *Para todo* $\mathbf{x} \in V_n(F)$ *y todo* $r \in F$, $r\mathbf{x} \in V_n(F)$.
S_2. *Para todo* $\mathbf{x} \in V_n(F)$, $1 \cdot \mathbf{x} = \mathbf{x}$.
S_3. *Para todo* $r, s \in F$ *y todo* $\mathbf{x} \in V_n(F)$, $r(s\mathbf{x}) = (rs)\mathbf{x}$.

S_4. *Para todo $r, s \in F$ y todo $\mathbf{x} \in V_n(F)$, $(r+s)\mathbf{x} = r\mathbf{x} + s\mathbf{x}$.*
S_5. *Para todo $r \in F$ y cualesquiera $\mathbf{x}$, $\mathbf{y} \in V_n(F)$, $r(\mathbf{x}+\mathbf{y}) = r\mathbf{x} + r\mathbf{y}$.*

PRUEBA. Véanse ejemplos 2.5, 2.6 y 2.7, pág. 17, y el problema 2, pág. 19.

En la notación de esta sección, el espacio vectorial n-dimensional V_n sobre el campo de los números reales que introdujimos en la sección 2 se denotaría por $V_n(R)$. Excepto en el capítulo 11, confinaremos nuestra atención a $V_n(R)$ y continuaremos usando para designarle la notación abreviada V_n. En el capítulo 11, tendremos ocasión de tratar con $V_2(C)$, el espacio bidimensional sobre el campo de los números complejos.

Un espacio vectorial V sobre un campo F puede definirse de modo abstracto como una estructura que tiene las propiedades enumeradas en el teorema 8.4. Cuando esto se hace, se introduce el concepto de independencia lineal y la dimensión del espacio vectorial puede definirse como el número máximo de vectores linealmente independientes en el espacio. Este número puede ser finito o puede no serlo. Si la dimensión es n (finita), entonces puede demostrarse que el espacio es esencialmente $V_n(F)$. Para detalles véase, por ejemplo, las referencias 22 o 27.

9. RESUMEN

En este capítulo introdujimos vectores n-dimensionales y consideramos las propiedades algebraicas de los vectores. Se introdujeron los conceptos de paralelismo y ortogonalidad entre vectores. Consideramos el producto escalar de dos vectores en el espacio n-dimensional y lo relacionamos con el concepto de ortogonalidad. Se discutieron los importantes conceptos de proyección ortogonal de un vector **a** sobre un vector no nulo **b** y componente de **a** en la dirección de **b**. La proyección ortogonal de **a** sobre **b** es un vector y el componente de **a** en la dirección de **b** es un número real - la longitud dirigida de la proyección de **a** sobre **b**. Aunque se dio una representación geométrica de los vectores, aún no se han usado éstos en el estudio de la geometría. En el próximo capítulo, estudiaremos geometría analítica tridimensional, a la que también llamaremos geometría analítica sólida.

Problemas de repaso

1. Sean $\mathbf{a} = (1, 2, 3, -2)$, $\mathbf{b} = (-3, 1, 6, 2)$ y $\mathbf{c} = (-2, 3, -1, 7)$. Encuéntrense

a) $\mathbf{a} + \mathbf{b}$ *b*) $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c})$
c) $\mathbf{a} \cdot \mathbf{a}$ *d*) $|\mathbf{b}|$
e) $\dfrac{\mathbf{c}}{|\mathbf{c}|}$ *f*) $2\mathbf{a} + \frac{1}{2}\mathbf{c}$.

2. ¿Cuáles de los siguientes pares de vectores están en la misma dirección?, ¿cuáles son paralelos?

a) $(2, 3), (-4, -6)$ *b*) $(1, -5), (-2, 15)$
c) $(3, -2), (4, -\frac{8}{3})$ *d*) $(2, 3, 5), (1, 2, 4)$.

3. Si $\mathbf{a} = (1, 5, 2, 4)$, $\mathbf{b} = (1, -2, 3, -1)$, calcúlese la longitud de

a) $\mathbf{a}$ *b*) $\mathbf{b}$
c) $\mathbf{a}+\mathbf{b}$ *b*) $\mathbf{a}-\mathbf{b}$.

4. Véanse si son o no ortogonales los siguientes pares de vectores.

a) $(1, 2), (-2, 1)$ *b*) $(1, 1, 1), (1, -1, 0)$
c) $(3, 5), (-3, 2)$ *d*) $(1, -2, 3, 5), (-1, 2, 0, 1)$.

5. Calcúlese en cada caso $\text{Comp}_{\mathbf{b}}\, \mathbf{a}$ y $\mathbf{Proy}_{\mathbf{b}}\, \mathbf{a}$.

a) $\mathbf{a} = (1, 1, 3)$, $\mathbf{b} = (-1, 2, 1)$
b) $\mathbf{a} = (1, 2, 1, 1)$, $\mathbf{b} = (-1, 3, -2, 2)$
c) $\mathbf{a} = (1, 1, -2)$, $\mathbf{b} = (3, -1, 1)$
d) $\mathbf{a} = (1, 5, 2)$, $\mathbf{b} = (-1, 0, 1)$.





Geometría analítica sólida

1. INTRODUCCIÓN

A mediados del siglo XIX el matemático irlandés William Rowan Hamilton (1805-1865) cambió su interés de la física matemática al álgebra y elaboró el álgebra de los números complejos basada en los pares ordenados de números reales. Después intentó desarrollar un álgebra de ternas y cuaternas de números. Uno de sus hijos, que sabía esto, le preguntó: "Bien, papá, ¿puedes multiplicar ternas?" Por lo que se dice contestó: "No, sólo puedo sumarlas y restarlas." Lo que sí descubrió fue un álgebra no conmutativa de dimensión cuatro (cuaternios).[1] La

[1] El problema de definir una multiplicación en V_n que dé a V_n una estructura de un álgebra con división tiene una historia larga e interesante. A mediados del siglo XIX, un matemático inglés, Arthur Cayley, mostró que esto era también posible para $n = 8$.

denominada parte pura del producto de cuaternios es, cuando se reduce a la dimensión tres, el "producto vectorial" que estudiaremos en este capítulo. Independientemente de Hamilton, el matemático alemán Hermann Gunther Grassmann (1809-1877) extendió este punto de vista de los números complejos a las n-adas ordenadas de números reales. Estos números hipercomplejos generalizaban los números complejos y los cuaternios de Hamilton. La contribución de Grassmann pasó inadvertida hasta su aplicación en 1915, en la teoría general de la relatividad, y es sólo hasta fecha muy reciente que su trabajo se ha apreciado plenamente. Fueron, sin embargo, dos físico-matemáticos los que se dieron cuenta de la significación que tenían los vectores en física, el norteamericano Josiah Willard Gibbs (1839-1903) y el inglés Oliver Heaviside (1850-1925), y el desarrollo del análisis vectorial tridimensional de la primera parte del presente siglo se debe en gran parte a estos dos hombres.

En este capítulo, el álgebra vectorial se aplica al estudio de la geometría euclidiana tridimensional y en la sección 12 algunos de los resultados obtenidos para el espacio tridimensional se generalizan para los espacios n-dimensionales. Las propiedades de los vectores bajo las operaciones de adición, multiplicación por un escalar y producto escalar que se obtuvieron en el capítulo 1 se usan en éste. Introducimos, además, en V_3 una nueva operación sobre vectores, el "producto vectorial". Este producto vectorial se aplica a dos vectores en V_3 y da como resultado un vector en V_3. El producto vectorial deriva su importancia del hecho de que tanto su longitud como su dirección tienen importante significación en geometría y física.

2. ESPACIO EUCLIDIANO TRIDIMENSIONAL

A fin de apreciar la terminología que va a introducirse en esta sección y también para ayudar a comprender la manera en que el álgebra vectorial se aplica en geometría, explicaremos primero cuáles son las imágenes geométricas que se encuentran tras el lenguaje que vamos a utilizar. Cuando denotamos un vector por una letra mayúscula —digamos $\mathbf{P} = (x, y, z)$—[1] indicamos que **P** ha de considerarse como un radio vector (es decir, el punto inicial de la flecha que representa **P** es el origen), o como el punto terminal de este radio vector (figura 1). Si $\mathbf{P} = (x, y, z)$ se llama punto, lo visualizamos como el punto terminal del radio vector $\mathbf{P} = (x, y, z)$, y los números x, y, z se llaman, entonces, *coordenadas* del punto **P**. Así pues, (x, y, z) puede llamarse un vector, un radio vector, o un punto, y los números x, y, z pueden llamarse componentes o coordenadas. El lenguaje que se utilice depende

Es, por tanto, posible para $n = 1, 2, 4,$ y 8, y solamente en fecha muy reciente (1958) se probó que éstas eran las únicas posibilidades.

[1] En la geometría tridimensional es una práctica común denotar las coordenadas por x, y, z en lugar de por x_1, x_2, x_3 y aquí seguiremos esa práctica.

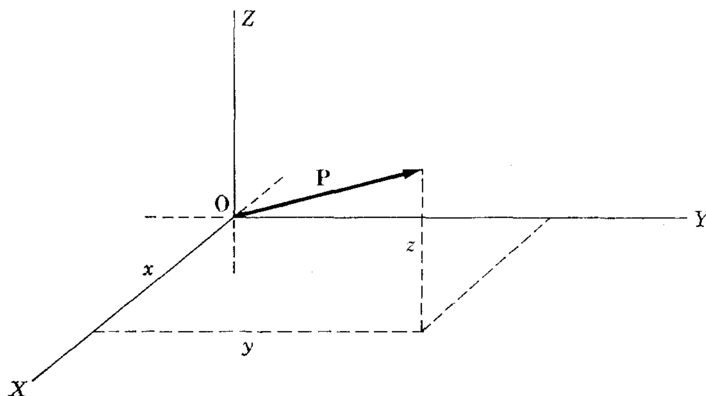


FIGURA 1 $(x,y,z) = \mathbf{P}$

de cual sea la aplicación que el usuario tiene in mente. Si denotamos un vector por una letra negrita $\mathbf{a} = (a_1, a_2, a_3)$, debemos entender que el vector se está usando para representar una dirección y una magnitud. En cualquier caso, los nombres que usemos no afectarán el álgebra de los vectores aunque la terminología puede resultar necesaria para comprender alguna aplicación particular del álgebra vectorial.

Ahora estamos preparados para dar una descripción de nuestro modelo analítico del espacio euclidiano. Llamamos a este modelo "espacio analítico euclidiano tridimensional" o, simplemente, "espacio euclidiano tridimensional". El espacio euclidiano tridimensional se denota por R^3 (léase: R tres). En la definición de R^3 debemos especificar qué es lo que vamos a entender por puntos, rectas y planos. Los puntos de R^3 son las ternas ordenadas (x, y, z) de V_3. Los números x, y, z han de visualizarse como las coordenadas rectangulares del punto $\mathbf{P} = (x, y, z)$ (figura 1). Nuestra definición de una recta en R^3 nace de la idea intuitiva de que una recta está determinada por un punto $\mathbf{P}_0$ y una dirección $\mathbf{a}$ ($\mathbf{a}$ es un vector no nulo) (figura 2). Los puntos $\mathbf{P}$ sobre la recta $\mathscr{L}$ que pasa por $\mathbf{P}_0$ en la dirección de $\mathbf{a}$ son, todos, puntos de la forma $\mathbf{P} = \mathbf{P}_0 + t\mathbf{a}$, donde t es un número real. Denotamos este conjunto por $\{\mathbf{P}_0 + t\mathbf{a} \mid t \in R\}$, lo que leemos: "el conjunto de todos los puntos $\mathbf{P}_0 + t\mathbf{a}$ con $t \in R$". Son todos los puntos que pueden alcanzarse desde $\mathbf{P}_0$, partiendo desde $\mathbf{P}_0$, siguiendo una dirección paralela a $\mathbf{a}$. La definición de un plano en R^3 surge de la idea de que un plano está determinado por dos rectas no paralelas $\mathscr{L}_1$ y $\mathscr{L}_2$ de direcciones respectivas $\mathbf{a}$ y $\mathbf{b}$ que se cortan en un punto $\mathbf{P}_0$ (figura 3). Los puntos $\mathbf{P}$ sobre el plano $\mathscr{P}$ determinado por $\mathscr{L}_1$ y $\mathscr{L}_2$ son, todos, los puntos de la forma $\mathbf{P} = \mathbf{P}_0 + u\mathbf{a} + v\mathbf{b}$, donde u y v son números reales. La distancia en el espacio euclidiano tridimensional R^3 desde un punto $\mathbf{P}_1$ a un punto $\mathbf{P}_2$ se define como la

longitud del vector $\mathbf{P}_2 - \mathbf{P}_1$ que va de $\mathbf{P}_1$ a $\mathbf{P}_2$. Nuestro espacio euclidiano tridimensional R^3 es, por tanto, V_3 con las nociones de punto, recta,

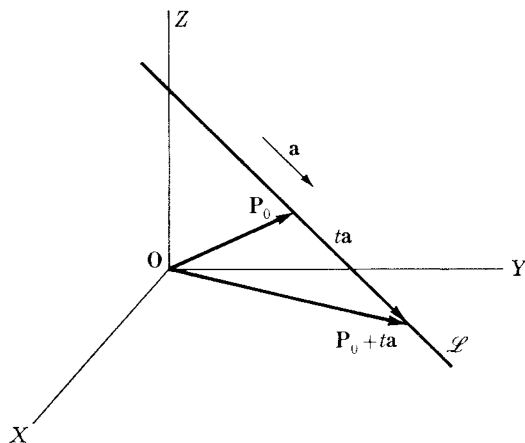


FIGURA 2

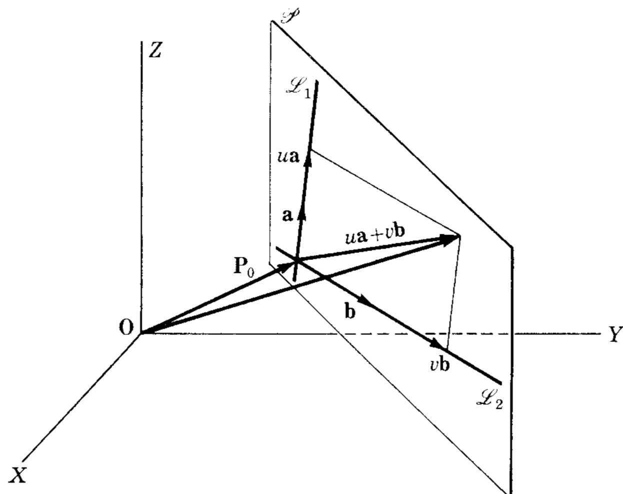


FIGURA 3 $\mathbf{P} = \mathbf{P}_0 + u\mathbf{a} + v\mathbf{b}$

plano y distancia definidas como acaba de explicarse. Damos ahora un enunciado preciso de la definición del espacio euclidiano tridimensional R^3.

2.1 Definición. *El* ***espacio*** *(analítico)* ***euclidiano tridimensional****, denotado por* R^3*, es el espacio vectorial tridimensional* V_3 *donde:*

1) *los elementos* (x, y, z) *de* V_3 *son los* ***puntos*** *de* R^3 *(figura 1);*

2) *un conjunto* $\mathscr{L}$ *de puntos de* R^3 *es una recta si hay un punto* $\mathbf{P}_0 \in \mathrm{R}^3$ *y un vector no nulo* $\mathbf{a} \in V_3$ *tal que (figura 2),*

$$\mathscr{L} = \{\mathbf{P}_0 + t\mathbf{a} \mid t \in \mathrm{R}\};$$

3) *un conjunto* $\mathscr{P}$ *de puntos de* R^3 *es un* ***plano*** *si hay un punto* $\mathbf{P}_0 \in \mathrm{R}^3$ *y dos vectores no paralelos* $\mathbf{a}$ *y* $\mathbf{b}$ *de* V_3 *tales que (figura 3),*

$$\mathscr{P} = \{\mathbf{P}_0 + u\mathbf{a} + v\mathbf{b} \mid u, v \in \mathrm{R}\};$$

4) *la* ***distancia****, denotada por* $d(\mathbf{P}_1, \mathbf{P}_2)$*, del punto* $\mathbf{P}_1 = (x_1, y_1, z_1)$ *al punto* $\mathbf{P}_2 = (x_2, y_2, z_2)$ *es la longitud del vector* $\mathbf{P}_2 - \mathbf{P}_1$*, es decir,*

$$d(\mathbf{P}_1, \mathbf{P}_2) = |\mathbf{P}_2 - \mathbf{P}_1| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

Las ecuaciones

$$\mathbf{P} = \mathbf{P}_0 + t\mathbf{a}$$

y

$$\mathbf{P} = \mathbf{P}_0 + u\mathbf{a} + v\mathbf{b}$$

se llaman *ecuaciones vectoriales* de la recta y el plano, respectivamente, y las ecuaciones correspondientes entre los componentes

$$x = x_0 + ta_1, \quad y = y_0 + ta_2, \quad z = z_0 + ta_3$$

y

$$x = x_0 + ua_1 + vb_1, \quad y = y_0 + ua_2 + vb_2, \quad z = z_0 + ua_3 + vb_3$$

se llaman *ecuaciones paramétricas* de la recta y el plano.

2.2 Ejemplo. Determínese una recta que pase poı los puntos $\mathbf{P}_0 = (1, 1, 2)$ y $\mathbf{P}_1 = (1, 2, 0)$.

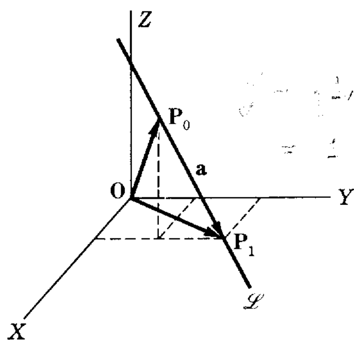


FIGURA 4

SOLUCIÓN. (Figura 4.) Sea $\mathbf{a} = \mathbf{P}_1 - \mathbf{P}_0 = (0, 1, -2)$. Entonces

2.3 $$\mathscr{L} = \{\mathbf{P}_0 + t(\mathbf{P}_1 - \mathbf{P}_0) \mid t \in \mathrm{R}\}$$

es claramente una recta que contiene a $\mathbf{P}_0$ y $\mathbf{P}_1$, ya que $\mathbf{P}_0$ resulta de hacer $t = 0$ y $\mathbf{P}_1$ de hacer $t = 1$. Luego

$$\mathscr{L} = \{(1, 1, 2) + t(0, 1, -2)\} = \{(1, 1+t, 2-2t)\}$$

es una recta que pasa por $(1, 1, 2)$ y $(1, 2, 0)$.

2.4 Ejemplo. Determínese un plano que pase por los puntos $\mathbf{P}_0 = (1, 0, 0)$, $\mathbf{P}_1 = (0, 1, 0)$ y $\mathbf{P}_2 = (0, 0, 1)$.

SOLUCIÓN. (Figura 5.) Definamos $\mathbf{a} = \mathbf{P}_1 - \mathbf{P}_0$, $\mathbf{b} = \mathbf{P}_2 - \mathbf{P}_0$, y

$$\begin{aligned}\mathscr{P} &= \{\mathbf{P}_0 + u\mathbf{a} + v\mathbf{b} \mid u, v \in \mathrm{R}\}\\ &= \{\mathbf{P}_0 + u(\mathbf{P}_1 - \mathbf{P}_0) + v(\mathbf{P}_2 - \mathbf{P}_0)\}.\end{aligned}$$

Como $\mathbf{a} = (-1, 1, 0)$ y $\mathbf{b} = (-1, 0, 1)$ no son paralelos, $\mathscr{P}$ es un plano que contiene a $\mathbf{P}_0$, $\mathbf{P}_1$, y $\mathbf{P}_2$; $\mathbf{P}_0$ resulta de hacer $u = 0$, $v = 0$; $\mathbf{P}_1$ resulta de hacer $u = 1$, $v = 0$, y $\mathbf{P}_2$ de hacer $u = 0$, $v = 1$. De donde resulta que

$$\begin{aligned}\mathscr{P} &= \{(1, 0, 0) + u(-1, 1, 0) + v(-1, 0, 1)\}\\ &= \{(1 - u - v, u, v)\}\end{aligned}$$

es un plano que pasa por los puntos $(1, 0, 0)$, $(0, 1, 0)$, y $(0, 0, 1)$.

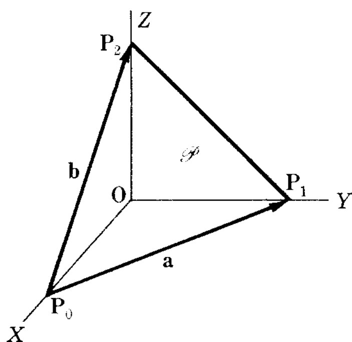


FIGURA 5

2.5 Ejemplo. Determínese la intersección de la recta $\mathscr{L}$ del ejemplo 2.2 con el plano $\mathscr{P}$ del ejemplo 2.4.

SOLUCIÓN. Supongamos que $\mathbf{P}$ es un punto de la intersección de $\mathscr{L}$ y $\mathscr{P}$. La intersección de dos conjuntos $\mathscr{L}$ y $\mathscr{P}$, denotada por $\mathscr{L} \cap \mathscr{P}$, es el conjunto

de todos los puntos que pertenecen tanto a $\mathscr{L}$ como a $\mathscr{P}$. Si $\mathbf{P} \in \mathscr{L} \cap \mathscr{P}$, entonces $\mathbf{P} \in \mathscr{L}$ de modo que

$$\mathbf{P} = (1, 1+t, 2-2t) \quad \text{para algún } t \in \mathrm{R}$$

y $\mathbf{P} \in \mathscr{P}$ de modo que

$$\mathbf{P} = (1-u-v, u, v) \quad \text{para ciertos } u, v \in \mathrm{R}.$$

Por tanto, $\mathbf{P}$ pertenece a la intersección si y sólo si

$$(1, 1+t, 2-2t) = (1-u-v, u, v) \quad \text{para algún } t, u, v \in \mathrm{R},$$

es decir, si y sólo si

$$\begin{aligned} 1 &= 1-u-v \\ 1+t &= u \\ 2-2t &= v. \end{aligned}$$

Resolviendo estas ecuaciones, encontramos

$$t = 3, \quad u = 4, \quad v = -4.$$

Por tanto,

$$\mathbf{P} = (1, 1+t, 2-2t) = (1, 4, -4) = (1-u-v, u, v).$$

Luego $(1, 4, -4)$ es *el* punto de intersección.

Problemas

1. Encuéntrese la distancia entre los siguientes pares de puntos de R^3:

a) $(1, 5, 3)$ y $(0, 0, 0)$
b) $(-2, 4, 3)$ y $(1, 8, -2)$
c) $(2, 1, -5)$ y $(-1, 0, 4)$
c) $(1, -9, 3)$ y $(-7, -2, 1)$
e) $(x_1, y_1, 0)$ y $(x_2, y_2, 0)$
f) $(0, 0, 0)$ y (x, y, z)
g) $\mathbf{P}_0$ y $\mathbf{P}_0 + t\mathbf{a}$
h) $\mathbf{P}_0$ y $(1-t)\mathbf{P}_0 + t\mathbf{P}_1$.

2. Determínese una recta que pasa por el punto $\mathbf{P}_0$ paralela a $\mathbf{a}$ cuando:

a) $\mathbf{P}_0 = (0, 0, 0)$ y $\mathbf{a} = (1, 1, 1)$
b) $\mathbf{P}_0 = (5, 3, -2)$ y $\mathbf{a} = (2, -3, 2)$
c) $\mathbf{P}_0 = (7, 12, -11)$ y $\mathbf{a} = (0, 0, 1)$
d) $\mathbf{P}_0 = (5, 7, 1)$ y $\mathbf{a} = (-2, 3, -5)$
e) $\mathbf{P}_0 = (-3, 2, -1)$ y $\mathbf{a} = (1, 5, -4)$
f) $\mathbf{P}_0 = (-1, -3, -5)$ y $\mathbf{a} = (-2, -7, -3)$.

3. Determínese en cada caso una recta que pase por los pares de puntos dados y proporciónese su ecuación paramétrica:

a) $(0, 0, 0)$ y $(1, 1, 1)$
b) $(1, 0, 0)$ y $(0, 1, 0)$
c) $(8, -3, 2)$ y $(5, 0, 0)$
d) $(5, 8, 1)$ y $(2, 6, -1)$
e) $(-3, 2, -1)$ y $(-2, 7, -5)$
f) $(1, 1, 1)$ y $(-3, 2, -1)$.

4. Determínese, en cada caso, un plano que pase por los puntos dados y proporciónese su ecuación paramétrica:

a) (2, 3, 1), (1, 1, −4) y (−3, 4, −2)
b) (1, 1, 0), (2, 0, 1) y (−1, 6, −1)
c) (1, 1, 1), (0, 0, 0) y (2, 0, 0)
d) (2, 3, 0), (−5, 1, 1) y (0, 1, 1)
e) (1, 1, 1), (3, −2, 0) y (4, 3, −1)
f) (0, 0, 0), (2, 3, −5) y (1, 2, 0).

5. ¿Son colineales los puntos de los siguientes conjuntos?

a) (0, 0, 0), (1, 1, 1), (−1, −1, −1)
b) (2, 3, −5), (0, 0, 0), (3, −2, 0)
c) (1, 2, 0), (5, −7, 8), (4, 3, −1).

6. ¿Cuál es una condición necesaria y suficiente para que tres puntos $\mathbf{P}_1$, $\mathbf{P}_2$ y $\mathbf{P}_3$ sean colineales?

7. Encuéntrese la intersección de la recta $\mathscr{L}$ y el plano $\mathscr{P}$ en cada uno de los siguientes casos:

a) $\mathscr{L} = \{(1, 1, 1)+t(2, 3, 4)\}$, $\mathscr{P} = \{(2, 3, 4)+u(1, 1, 1)+v(1, 0, -2)\}$
b) $\mathscr{L} = \{(1, 2, 0)+t(-5, 1, 1)\}$, $\mathscr{P} = \{(2, 3, 1)+u(2, 0, 0)+v(2, 6, -1)\}$
c) $\mathscr{L}$ pasa por (0, 0, 0) y (1, 1, 1), y $\mathscr{P}$ pasa por (2, 3, 1), (1, 1, −4) y (−3, 4, 2)
d) $\mathscr{L}$ pasa por (8, −3, 2) y (5, 0, 0), y $\mathscr{P}$ pasa por (1, 1, 1), (0, 0, 0), y (2, 0, 0)
e) $\mathscr{L}$ pasa por (−3, 2, −1) y (−2, 7, −5), y $\mathscr{P}$ pasa por (1, 1, 1), (3, −2, 0) y (4, 3, −1)
f) $\mathscr{L}$ pasa por (1, 1, 1) y (−3, 2, −1), y $\mathscr{P}$ pasa por (2, 3, 1), (1, 1, −4) y (−3, 4, 2).

3. RECTAS

En esta sección demostraremos que dos puntos distintos determinan inequívocamente una recta, es decir, que hay una recta y sólo una recta que pasa por cada par de puntos distintos de R^3. Antes de probar este resultado daremos un breve repaso de las operaciones entre conjuntos y su notación.

Un conjunto $\mathscr{A}$ se dice que es un subconjunto de un conjunto $\mathscr{B}$, y entonces se escribe $\mathscr{A} \subset \mathscr{B}$, si todo elemento de $\mathscr{A}$ es también un elemento de $\mathscr{B}$. Decimos que dos conjuntos son iguales si y sólo si son idénticos. Así pues, $\mathscr{A} = \mathscr{B}$ si y sólo si $\mathscr{A} \subset \mathscr{B}$ y $\mathscr{B} \subset \mathscr{A}$. Hemos tenido ya ocasión, en el ejemplo 2.5, de considerar la intersección de conjuntos. La *intersección* de los conjuntos $\mathscr{A}$ y $\mathscr{B}$, escrita $\mathscr{A} \cap \mathscr{B}$, es el conjunto de todos los elementos

que se encuentran tanto en $\mathscr{A}$ como en $\mathscr{B}$, es decir, en los dos a la vez. Para que la intersección de dos conjuntos sea siempre un conjunto, es conveniente introducir el *conjunto nulo* o *conjunto vacío*. El conjunto vacío es el conjunto que no tiene elementos y se denota por $\varnothing$. El conjunto vacío es un subconjunto de todo conjunto. La *unión* de los conjuntos $\mathscr{A}$ y $\mathscr{B}$, escrita $\mathscr{A} \cup \mathscr{B}$, es el conjunto de todos los elementos que están en $\mathscr{A}$ o en $\mathscr{B}$.

3.1 Teorema. *Para cada par de puntos distintos de* R^3 *hay una y sólo una recta que pasa por ellos.*

PRUEBA. Sean $\mathbf{P}_1$ y $\mathbf{P}_2$ un par de puntos distintos de R^3. Entonces

$$\mathscr{L} = \{\mathbf{P}_1 + t(\mathbf{P}_2 - \mathbf{P}_1) \mid t \in \mathrm{R}\}$$

es una recta que pasa por $\mathbf{P}_1$ y $\mathbf{P}_2$. Supongamos que

$$\mathscr{L}' = \{\mathbf{P}_0 + s\mathbf{a} \mid s \in \mathrm{R}\}$$

es una recta que pasa por $\mathbf{P}_1$ y $\mathbf{P}_2$. Deseamos demostrar que $\mathscr{L}' = \mathscr{L}$. Como $\mathbf{P}_1$ y $\mathbf{P}_2$ son puntos en $\mathscr{L}'$ existen números reales, distintos, s_1 y s_2 tales que $\mathbf{P}_1 = \mathbf{P}_0 + s_1\mathbf{a}$ y $\mathbf{P}_2 = \mathbf{P}_0 + s_2\mathbf{a}$. Si $\mathbf{P} \in \mathscr{L}$, entonces, para algún $t \in \mathrm{R}$

$$\begin{aligned}\mathbf{P} &= \mathbf{P}_1 + t(\mathbf{P}_2 - \mathbf{P}_1) = \mathbf{P}_0 + s_1\mathbf{a} + t(s_2 - s_1)\mathbf{a} \\ &= \mathbf{P}_0 + [s_1 + t(s_2 - s_1)]\mathbf{a}.\end{aligned}$$

Así pues, $\mathbf{P} \in \mathscr{L}'$ y $\mathscr{L} \subset \mathscr{L}'$. Recíprocamente, si $\mathbf{P} \subset \mathscr{L}'$, entonces, para algún $s \in \mathrm{R}$

$$\mathbf{P} = \mathbf{P}_0 + s\mathbf{a} = \mathbf{P}_1 - s_1\mathbf{a} + s\mathbf{a} = \mathbf{P}_1 + \frac{s - s_1}{s_2 - s_1}(\mathbf{P}_2 - \mathbf{P}_1).$$

Luego $\mathbf{P} \in \mathscr{L}$ y $\mathscr{L}' \subset \mathscr{L}$. Como $\mathscr{L}' \subset \mathscr{L}$ y $\mathscr{L} \subset \mathscr{L}'$, tenemos $\mathscr{L}' = \mathscr{L}$.

3.2 Definición. *Dos rectas*

$$\mathscr{L}_1 = \{\mathbf{P}_1 + s\mathbf{a} \mid s \in \mathrm{R}\} \quad y \quad \mathscr{L}_2 = \{\mathbf{P}_2 + t\mathbf{b} \mid t \in \mathrm{R}\}$$

se dicen ***paralelas*** *si los vectores* $\mathbf{a}$ *y* $\mathbf{b}$ *son paralelos.*

3.3 Corolario. *Para todo punto* $\mathbf{P}_1 \in \mathrm{R}^3$ *y toda recta* $\mathscr{L} = \{\mathbf{P}_0 + s\mathbf{a} \mid s \in \mathrm{R}\}$, *hay una y solamente una recta que pasa por* $\mathbf{P}_1$ *paralela a* $\mathscr{L}$.

PRUEBA. $\mathscr{L}_1 = \{\mathbf{P}_1 + t\mathbf{a} \mid t \in \mathrm{R}\}$ es una recta que pasa por $\mathbf{P}_1$ y es paralela a $\mathscr{L}$. Sea $\mathscr{L}_2 = \{\mathbf{P}_2 + u\mathbf{b} \mid u \in \mathrm{R}\}$ otra recta que pasa por $\mathbf{P}_1$ y es paralela a $\mathscr{L}$. Como $\mathbf{P}_1 \in \mathscr{L}_2$, existe un número real u_1 tal que

$$\mathbf{P}_1 = \mathbf{P}_2 + u_1\mathbf{b}.$$

Además, $\mathscr{L}_2$ paralela a $\mathscr{L}$ implica que $\mathbf{a} = r\mathbf{b}$ para algún $r \in \mathrm{R}$. De donde

$$\mathbf{P}_1 + \mathbf{a} = \mathbf{P}_2 + (u_1 + r)\mathbf{b}$$

y $\mathbf{P}_1 + \mathbf{a} \in \mathscr{L}_1 \cap \mathscr{L}_2$. Como $\mathbf{P}_1$ y $\mathbf{P}_1 + \mathbf{a}$ son puntos distintos de $\mathscr{L}_1 \cap \mathscr{L}_2$, de acuerdo con el teorema 3.1, $\mathscr{L}_1 = \mathscr{L}_2$.

3.4 Corolario. *Si* $\mathscr{L}_1 = \{\mathbf{P}_1 + s\mathbf{a} \mid s \in \mathrm{R}\}$ y $\mathscr{L}_2 = \{\mathbf{P}_2 + t\mathbf{b} \mid t \in \mathrm{R}\}$ *son rectas paralelas, entonces* $\mathscr{L}_1 = \mathscr{L}_2$ *o* $\mathscr{L}_1 \cap \mathscr{L}_2 = \varnothing$.

PRUEBA. Supongamos que $\mathscr{L}_1 \cap \mathscr{L}_2 \neq \varnothing$ y sea $\mathbf{P}_0$ un punto de $\mathscr{L}_1 \cap \mathscr{L}_2$. Entonces existen números reales s_0 y t_0 tales que

$$\mathbf{P}_0 = \mathbf{P}_1 + s_0\mathbf{a} = \mathbf{P}_2 + t_0\mathbf{b}.$$

Además, como $\mathscr{L}_1$ y $\mathscr{L}_2$ son paralelas, $\mathbf{a} = r\mathbf{b}$ para algún $r \in \mathrm{R}$. De donde

$$\mathbf{P}_0 + \mathbf{a} = \mathbf{P}_1 + (s_0 + 1)\mathbf{a} = \mathbf{P}_2 + (t_0 + r)\mathbf{b}$$

y $\mathbf{P}_0 + \mathbf{a} \in \mathscr{L}_1 \cap \mathscr{L}_2$. Como $\mathbf{P}_0$ y $\mathbf{P}_0 + \mathbf{a}$ son puntos distintos en $\mathscr{L}_1$ y en $\mathscr{L}_2$, según el teorema 3.1, $\mathscr{L}_1 = \mathscr{L}_2$.

3.5 Corolario. *Si las rectas* $\mathscr{L}_1$ *y* $\mathscr{L}_2$ *no son paralelas, entonces* $\mathscr{L}_1 \cap \mathscr{L}_2$ *es vacío o consiste en un solo punto.*

PRUEBA. Si $\mathscr{L}_1 \cap \mathscr{L}_2$ contiene más de un punto, entonces, de acuerdo con el teorema 3.1, tendríamos $\mathscr{L}_1 = \mathscr{L}_2$. Sin embargo $\mathscr{L}_1$ y $\mathscr{L}_2$ no son paralelas y no pueden, por tanto, coincidir. De esta manera, $\mathscr{L}_1 \cap \mathscr{L}_2$ no puede contener más de un punto.

3.6 Ejemplo. Determínese si las siguientes rectas son paralelas o no lo son y dígase cual es su intersección:

$$\mathscr{L}_1 = \{(1, 3, -2) + t(3, -6, 9)\}, \quad \mathscr{L}_2 = \{(2, 1, 7) + s(-2, 4, -6)\}.$$

SOLUCIÓN. Las rectas $\mathscr{L}_1$ y $\mathscr{L}_2$ son paralelas si hay un número real r tal que

$$(3, -6, 9) = r(-2, 4, -6).$$

Es decir, si $3 = -2r$, $-6 = 4r$ y $9 = -6r$. Estas ecuaciones se satisfacen por $r = -\frac{3}{2}$:

$$(3, -6{,}9) = -\tfrac{3}{2}(-2, 4, -6)$$

y las rectas son paralelas. Como las rectas son paralelas, según el corolario 3.4 $\mathscr{L}_1 = \mathscr{L}_2$ o $\mathscr{L}_1 \cap \mathscr{L}_2 = \varnothing$. El punto $\mathbf{P}_1 = (1, 3, -2) \in \mathscr{L}_1$. Si $\mathbf{P}_1 \in \mathscr{L}_2$, entonces $\mathscr{L}_1 = \mathscr{L}_2$ y si $\mathbf{P}_1 \notin \mathscr{L}_2$,[1] entonces $\mathscr{L}_1 \cap \mathscr{L}_2 = \varnothing$.

[1] La notación $\mathbf{P}_1 \notin \mathscr{L}_2$ denota que $\mathbf{P}_1$ no es un elemento de $\mathscr{L}_2$.

Si $\mathbf{P}_1 \in \mathscr{L}_2$, entonces para algún $s \in \mathrm{R}$,

$$(1, 3, -2) = (2, 1, 7) + s(-2, 4, -6)$$

o

$$(1, 3, -2) - (2, 1, 7) = (-1, 2, -9) = s(-2, 4, -6).$$

Esta última ecuación es equivalente a las tres ecuaciones componentes

$$-1 = -2s, \quad 2 = 4s, \quad -9 = -6s.$$

Pero no hay ningún número s que satisfaga simultáneamente estas tres ecuaciones y, por tanto, $\mathbf{P}_1 \notin \mathscr{L}_2$. Así que $\mathscr{L}_1 \cap \mathscr{L}_2 = \varnothing$.

3.7 Ejemplo. Determínese si los siguientes pares de rectas son o no paralelos y determínese su intersección:

$$\mathscr{L}_1 = \{(1, 3, -2) + t(3, -6, 9)\}, \quad \mathscr{L}_2 = \{(2, 1, 7) + s(1, -3, 4)\}.$$

SOLUCIÓN. Las rectas $\mathscr{L}_1$ y $\mathscr{L}_2$ son paralelas si para algún $r \in \mathrm{R}$

$$(3, -6, 9) = r(1, -3, 4).$$

Como no hay ningún número r que satisfaga esta ecuación, $\mathscr{L}_1$ y $\mathscr{L}_2$ no son paralelas. Luego $\mathscr{L}_1 \cap \mathscr{L}_2 = \varnothing$ o $\mathscr{L}_1 \cap \mathscr{L}_2$ contiene un punto. Si $\mathscr{L}_1 \cap \mathscr{L}_2 \neq \varnothing$ hay un punto $\mathbf{P}_0 \in \mathscr{L}_1 \cap \mathscr{L}_2$. Es decir, hay números $t, s \in \mathrm{R}$ tales que

$$\mathbf{P}_0 = (1, 3, -2) + t(3, -6, 9) = (2, 1, 7) + s(1, -3, 4)$$

o

$$t(3, -6, 9) - s(1, -3, 4) = (2, 1, 7) - (1, 3, -2) = (1, -2, 9).$$

Esta ecuación es equivalente a las tres ecuaciones de componentes

$$\begin{aligned} 3t - s &= 1 \\ -6t + 3s &= -2 \\ 9t - 4s &= 9. \end{aligned}$$

Resolviendo las dos primeras ecuaciones para s y t, encontramos $s = 0$ y $t = \frac{1}{3}$. Como estos valores no satisfacen la tercera ecuación, no hay ninguna solución para el sistema y $\mathscr{L}_1 \cap \mathscr{L}_2 = \varnothing$.

3.8 Definición. *θ es un* ***ángulo*** *entre las rectas $\mathscr{L}_1$ y $\mathscr{L}_2$ si para ciertos vectores no nulos* $\mathbf{a}$ *y* $\mathbf{b}$, $\mathscr{L}_1 = \{\mathbf{P}_1 + s\mathbf{a}\}$, $\mathscr{L}_2 = \{\mathbf{P}_2 + t\mathbf{b}\}$, *y θ es el ángulo entre* $\mathbf{a}$ *y* $\mathbf{b}$.

Hablaremos de ángulo entre dos rectas aun en el caso de que no se intersecten.

3.9 Ejemplo. Encuéntrese un ángulo entre las dos rectas del ejemplo 3.7.

SOLUCIÓN. Sean $\mathbf{a} = (3, -6, 9)$ y $\mathbf{b} = (1, -3, 4)$. Si θ es el ángulo entre $\mathbf{a}$ y $\mathbf{b}$, entonces

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{3+18+36}{3\sqrt{14}\sqrt{26}} = \frac{57}{6\sqrt{91}} = 0.99587$$

y θ tiene como medida en grados $5°12'$ o $354°48'$.

A veces es conveniente expresar los vectores de V_3 en términos de los vectores unitarios (figura 6)

3.10 $$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1).$$

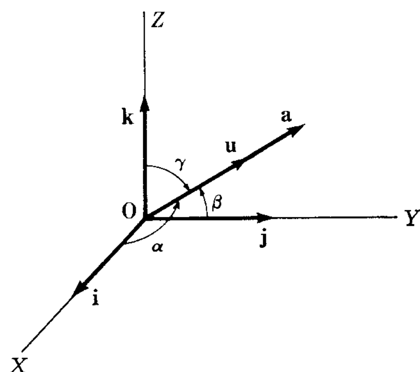


FIGURA 6 $\mathbf{u} = (\cos\alpha, \cos\beta, \cos\gamma)$

Estos vectores apuntan en las direcciones positivas de los ejes X, Y y Z, respectivamente. Para cualquier vector $\mathbf{a}\in V_3$, tenemos

$$\begin{aligned}\mathbf{a} &= (a_1, a_2, a_3) = (a_1, 0, 0)+(0, a_2, 0)+(0, 0, a_3)\\ &= a_1(1, 0, 0)+a_2(0, 1, 0)+a_3(0, 0, 1)\end{aligned}$$

o bien,

3.11 $$\mathbf{a} = a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}.$$

Sea $\mathbf{a} = (l, m, n)$ un vector no nulo paralelo a una recta $\mathscr{L}$. Los números l, m, n se llaman *números directores* de la recta $\mathscr{L}$. Sea α el ángulo entre $\mathbf{i}$ y $\mathbf{a}$; β el ángulo entre $\mathbf{j}$ y $\mathbf{a}$, y γ el ángulo entre $\mathbf{k}$ y $\mathbf{a}$ (figura 6). Los ángulos α, β, y γ se llaman *ángulos directores* de $\mathscr{L}$, y $\cos\alpha$, $\cos\beta$, $\cos\gamma$ se llaman *cosenos directores* de $\mathscr{L}$. Si $\mathbf{u}$ es el vector unitario en la dirección de $\mathbf{a}$:

$$\mathbf{u} = \frac{\mathbf{a}}{|\mathbf{a}|} = \frac{(l, m, n)}{\sqrt{l^2+m^2+n^2}},$$

de donde se muestra fácilmente que (problema 4)

3.12 $$\cos\alpha = \mathbf{i}\cdot\mathbf{u}, \quad \cos\beta = \mathbf{j}\cdot\mathbf{u}, \quad \cos\gamma = \mathbf{k}\cdot\mathbf{u}$$

3.13 $$\mathbf{u} = (\cos\alpha, \cos\beta, \cos\gamma)$$

3.14 $$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1 .$$

Sean α_1, β_1, γ_1 y α_2, β_2, γ_2 los ángulos directores de las rectas $\mathscr{L}_1$ y $\mathscr{L}_2$, respectivamente. Si los ángulos de dirección de $\mathscr{L}_1$ y $\mathscr{L}_2$ están determinados por los vectores $\mathbf{a}_1$ y $\mathbf{a}_2$, respectivamente, y θ es el ángulo entre $\mathbf{a}_1$ y $\mathbf{a}_2$, entonces

3.15 $$\cos\theta = \cos\alpha_1\cos\alpha_2 + \cos\beta_1\cos\beta_2 + \cos\gamma_1\cos\gamma_2 .$$

Problemas

1. Sean

$$\begin{aligned}
\mathscr{L}_1 &= \{(2, 1, 4)+r(1, 1, 1)\}\\
\mathscr{L}_2 &= \{(1, -1, 4)+s(2, -1, 3)\}\\
\mathscr{L}_3 &= \{(1, -2, 5)+t(1, -3, 4)\}\\
\mathscr{L}_4 &= \{(3, -2, 7)+u(6, -3, 9)\}\\
\mathscr{L}_5 &= \{(3, 2, 3)+v(-2, -2, -2)\}
\end{aligned}$$

Determínese si son o no paralelos cada uno de los siguientes pares de rectas y determínense sus intersecciones.

a) $\mathscr{L}_1$ y $\mathscr{L}_2$ *b*) $\mathscr{L}_1$ y $\mathscr{L}_3$ *c*) $\mathscr{L}_1$ y $\mathscr{L}_4$
d) $\mathscr{L}_1$ y $\mathscr{L}_5$ *e*) $\mathscr{L}_2$ y $\mathscr{L}_3$ *f*) $\mathscr{L}_2$ y $\mathscr{L}_4$.

2. Identifíquese el conjunto de todos los puntos $\mathbf{P} = (x, y, z)$ que satisfacen

$$\frac{x-3}{2} = \frac{y+5}{3} = \frac{z-4}{-2} .$$

3. Determínense:

a) Los ángulos entre una recta paralela al vector (1, 1, 1) y los ejes de coordenadas;

b) los ángulos entre la recta que pasa por los puntos (1, 0, 1) y (0, 1, 0) y los ejes de coordenadas;

c) un ángulo entre las rectas de los incisos (*a*) y (*b*).

4. Demuéstrese que las ecuaciones 3.12, 3.13, 3.14 y 3.15 se verifican.

5. Determínense:

a) las rectas que pasan por el origen con ángulos directores $\alpha = 45°$, $\beta = 60°$;

b) las rectas que pasan por el punto $(-2, 7, 13)$ con ángulos directores $\alpha = \beta = 45^\circ$;

c) las rectas que pasan por el origen con ángulos directores $\alpha = \beta = \gamma$.

6. Demuéstrese que los vectores unitarios **i**, **j**, **k** definidos por 3.10 satisfacen las relaciones

a) $\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1$

b) $\mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{i} = 0$

4. EL PRODUCTO VECTORIAL

El plano $\mathscr{P} = \{\mathbf{P}_0 + u\mathbf{a} + v\mathbf{b} \mid u, v \in \mathrm{R}\}$ puede describirse como el conjunto de todos los puntos **P** tales que $\mathbf{P}-\mathbf{P}_0$ es ortogonal a un vector **n** donde **n** es ortogonal tanto a **a** como a **b**. Esto será demostrado en la sección 7. En esta sección demostramos la manera en que un vector **n** puede encontrarse dados los vectores **a** y **b**. Para este fin introducimos una operación sobre vectores de V_3 a la que llamamos "producto vectorial". Cuando se aplica a vectores **a** y **b** el espacio vectorial nos da como resultado un vector ortogonal tanto a **a** como a **b**. Aparte de este uso geométrico en la descripción de un plano, el producto vectorial tiene otras importantes aplicaciones en geometría y en física.

4.1 Definición. *El **producto vectorial** de dos vectores* $\mathbf{a} = (a_1, a_2, a_3)$ *y* $\mathbf{b} = (b_1, b_2, b_3)$ *de* V_3, *denotado por* $\mathbf{a}\times\mathbf{b}$, *lo que leeremos* "**a** *cruz* **b**", *es el vector definido por*

$$\mathbf{a}\times\mathbf{b} = (a_2b_3 - a_3b_2,\ a_3b_1 - a_1b_3,\ a_1b_2 - a_2b_1).$$

El producto $\mathbf{a}\times\mathbf{b}$ es un vector, y, como

$$\mathbf{a}\cdot(\mathbf{a}\times\mathbf{b}) = a_1(a_2b_3 - a_3b_2) + a_2(a_3b_1 - a_1b_3) + a_3(a_1b_2 - a_2b_1) = 0$$

y

$$\mathbf{b}\cdot(\mathbf{a}\times\mathbf{b}) = b_1(a_2b_3 - a_3b_2) + b_2(a_3b_1 - a_1b_3) + b_3(a_1b_2 - a_2b_1) = 0,$$

$\mathbf{a}\times\mathbf{b}$ es ortogonal tanto a **a** como a **b**.

Las propiedades fundamentales del producto vectorial son:
Para **a**, **b**, $\mathbf{c}\in V_3$ *cualesquiera y todo* $r\in\mathrm{R}$,

4.2 $$\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a}$$

4.3 $$(r\mathbf{a})\times\mathbf{b} = r(\mathbf{a}\times\mathbf{b})$$

4.4 $$\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = \mathbf{a}\times\mathbf{b} + \mathbf{a}\times\mathbf{c}.$$

La ecuación 4.2 nos dice que el producto vectorial no es conmutativo (es anticonmutativo); la ecuación 4.3 muestra la relación entre la multi-

plicación por un número real y el producto vectorial; y la ecuación 4.4 nos dice que el producto vectorial es distributivo respecto a la adición. Estas propiedades son simples consecuencias de la definición 4.1 y las propiedades de los números reales (problema 5). Es fácil construir ejemplos que nos muestran que, en general, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$, es decir, que la ley asociativa no se verifica (problema 6).

Los vectores unitarios $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, y $\mathbf{k} = (0, 0, 1)$ satisfacen las relaciones

4.5
$$\begin{aligned} \mathbf{i} \times \mathbf{i} &= \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0} \\ \mathbf{i} \times \mathbf{j} &= \mathbf{k} = -\mathbf{j} \times \mathbf{i} \\ \mathbf{j} \times \mathbf{k} &= \mathbf{i} = -\mathbf{k} \times \mathbf{j} \\ \mathbf{k} \times \mathbf{i} &= \mathbf{j} = -\mathbf{i} \times \mathbf{k}. \end{aligned}$$

Las ecuaciones 4.5 son fáciles de recordar. Los productos $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$ y $\mathbf{k} \times \mathbf{i} = \mathbf{j}$ se corresponden con las permutaciones cíclicas de $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, a saber, $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, $\{\mathbf{j}, \mathbf{k}, \mathbf{i}\}$, y $\{\mathbf{k}, \mathbf{i}, \mathbf{j}\}$. Usando las propiedades 4.3 y 4.4 y los resultados de 4.5, podemos obtener el producto vectorial de dos vectores cualesquiera de V_3 como sigue:

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= (a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}) \times (b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k}) \\ &= a_1 b_1(\mathbf{i}\times\mathbf{i})+a_1 b_2(\mathbf{i}\times\mathbf{j})+a_1 b_3(\mathbf{i}\times\mathbf{k})+a_2 b_1(\mathbf{j}\times\mathbf{i})+a_2 b_2(\mathbf{j}\times\mathbf{j}) \\ &\quad +a_2 b_3(\mathbf{j}\times\mathbf{k})+a_3 b_1(\mathbf{k}\times\mathbf{i})+a_3 b_2(\mathbf{k}\times\mathbf{j})+a_3 b_3(\mathbf{k}\times\mathbf{k}) \\ &= (a_2 b_3-a_3 b_2)\mathbf{i}+(a_3 b_1-a_1 b_3)\mathbf{j}+(a_1 b_2-a_2 b_1)\mathbf{k}. \end{aligned}$$

Una representación más conveniente del producto vectorial puede darse en términos de determinantes. Una matriz $m \times n$ de números reales A es una función que tiene como dominio el conjunto de pares de enteros $\{(i,j) \mid 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n\}$ y el rango en R. Un valor de la función $A(i,j)$ que se representa por a_{ij} y que se llama entrada, y la matriz se describe desplegando las entradas en forma rectangular. Asociamos con cada matriz cuadrada ($m = n$) un número al que llamamos determinante de la matriz. El determinante de una matriz 2×2 (2 renglones horizontales y 2 columnas verticales) se define como sigue:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22}-a_{12}a_{21}.$$

El determinante de una matrix 3×3

4.6
$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

puede definirse en términos de matrices 2×2. Los determinantes

$$\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

se llaman *menores* de las entradas a_{11}, a_{12}, a_{13}, respectivamente, del primer renglón. El menor de a_{ij} (el primer subíndice indica el renglón en que a_{ij} se encuentra y el segundo la columna) es el determinante que se obtiene omitiendo el renglón i-ésimo y la columna j-ésima en la matriz 4.6, es decir, tachando el renglón y la columna en las que a_{ij} se encuentran y formando el determinante de la restante matriz 2×2. El determinante de la matriz 3×3 que aparece en 4.6, se define como

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Esto se llama "desarrollo del determinante en términos de los menores del primer renglón". Por tanto

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Puede asignarse un significado al determinante si los números del primer renglón se reemplazan por vectores. Escribimos

4.7
$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \mathbf{i}(a_2 b_3 - a_3 b_2) + \mathbf{j}(a_3 b_1 - a_1 b_3) + \mathbf{k}(a_1 b_2 - a_2 b_1).$$

Se ha señalado anteriormente que el producto vectorial $\mathbf{a} \times \mathbf{b}$ es ortogonal tanto a $\mathbf{a}$ como a $\mathbf{b}$. La longitud del vector $\mathbf{a} \times \mathbf{b}$ tiene un significado geométrico. Calculando el cuadrado de la longitud de $\mathbf{a} \times \mathbf{b}$ obtenemos, por el álgebra elemental,

$$\begin{aligned} |\mathbf{a} \times \mathbf{b}|^2 &= (a_2 b_3 - a_3 b_2)^2 + (a_3 b_1 - a_1 b_3)^2 + (a_1 b_2 - a_2 b_1)^2 \\ &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1 b_1 + a_2 b_2 + a_3 b_3)^2 ; \end{aligned}$$

es decir,

4.8
$$|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2 .$$

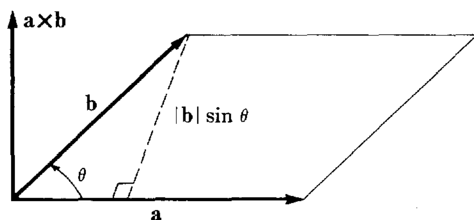


FIGURA 7

La ecuación 7.5 (pág. 69) afirma que

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos \theta,$$

donde θ es un ángulo entre **a** y **b**. Así pues, si tomamos θ como el ángulo $0 \leqslant \theta \leqslant \pi$ entre **a** y **b** (figura 7), tenemos

$$|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2\,|\mathbf{b}|^2\,(1 - \cos^2 \theta) = |\mathbf{a}|^2\,|\mathbf{b}|^2 \operatorname{sen}^2 \theta$$

y

4.9 $$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}| \operatorname{sen} \theta,$$

donde $\operatorname{sen} \theta \geqslant 0$ ya que $0 \leqslant \theta \leqslant \pi$. Como $|\mathbf{b}| \operatorname{sen} \theta$ es la altura del paralelogramo de lados **a** y **b** (figura 7), hemos demostrado que *la longitud de* **a** × **b** *es el área del paralelogramo de lados* **a** *y* **b**.

La ecuación 4.9 sugiere el siguiente teorema.

4.10 Teorema. *Dos vectores* **a**, **b** $\in V_3$ *son paralelos si y sólo si* $\mathbf{a} \times \mathbf{b} = 0$.

PRUEBA. $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ si y sólo si $|\mathbf{a} \times \mathbf{b}|^2 = 0$. De acuerdo con la ecuación 4.8 vemos que esto es equivalente a $(\mathbf{a} \cdot \mathbf{b})^2 = |\mathbf{a}|^2\,|\mathbf{b}|^2$ o $|\mathbf{a} \cdot \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|$. Esta última igualdad es la desigualdad de Schwarz (teorema 7.6, pág. 34) en el caso en que la igualdad se verifica, y ya hemos probado que en la desigualdad de Schwarz sólo se verifica la igualdad si **a** y **b** son paralelos.

Problemas

1. Sean $\mathbf{a} = (5, -2, 4)$, $\mathbf{b} = (2, 11, -1)$, $\mathbf{c} = (-7, 6, 9)$. Calcúlense:

a) $\mathbf{a} \times \mathbf{b}$
b) $\mathbf{b} \times \mathbf{a}$
c) $\mathbf{a} \times \mathbf{c}$
d) $\mathbf{a} \times (\mathbf{b} + \mathbf{c})$
e) $(2\mathbf{a}) \times (3\mathbf{b})$
f) $|\mathbf{a} \times \mathbf{b}|$
g) $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$
h) $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{c})$
i) $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b})$
j) $\mathbf{a} \times (\mathbf{a} \times \mathbf{b})$
k) $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$
l) $(\mathbf{a} + \mathbf{b}) \times (\mathbf{a} - \mathbf{c})$.

2. Establézcase la identidad

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{a}\cdot\mathbf{b})\mathbf{c}.$$

3. Usando la identidad del problema 2 y el teorema 4.10, pruébese que: **a** ortogonal tanto a **b** como a **c** implica **a** paralelo a $\mathbf{b}\times\mathbf{c}$.

4. Determínense todos los vectores no nulos ortogonales a:

a) (1, 0, 0) y (0, 1, 0) *b*) (1, 1, 1) y (1, 0, 0)
c) (2, −3, 4) y (−1, 5, 7) *d*) (1, −2, −4) y (−3, 2, −6)
e) (2, 6, −4) y (3, 9, −6) *f*) (−1, 1, 2) y (1, 1, 1).

5. Pruébense: *a*) 4.2 *b*) 4.3 *c*) 4.4.

6. Usando la ecuación 4.5 (pág. 55) encuéntrense $\mathbf{i}\times(\mathbf{i}\times\mathbf{j})$ e $(\mathbf{i}\times\mathbf{i})\times\mathbf{j}$ para demostrar con ello que la ley asociativa no se verifica para el producto vectorial.

7. Demuéstrese que:

a)
$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix}$$

b)
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$$

8. Demuéstrese que: $\mathbf{a} = (a_1, a_2)$ y $\mathbf{b} = (b_1, b_2)$ son paralelos si y sólo si

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1 = 0 .$$

9. Calcúlese el área de los paralelogramos de lados:

a) (5, 3, 0) y (3, 7, 0)
b) $\mathbf{i}-\mathbf{j}+5\mathbf{k}$ y $2\mathbf{i}+4\mathbf{j}-8\mathbf{k}$
c) (4, 13, −11) y (8, −10, 21)
d) (1, 3, 7) y (−2, −4, 3)
e) $2\mathbf{i}+3\mathbf{j}+5\mathbf{k}$ e $\mathbf{i}-2\mathbf{k}$
f) (−3, 2, −4) y (1, 1, 1).

10. Calcúlese el área de los triángulos con vértices:

a) (0, 0, 0), (1, 0, 0), (3, 8, 0)
b) (5, 0, 16), (8, 4, 12), (1, −1, 1)
c) $5\mathbf{i}-4\mathbf{j}$, $12\mathbf{k}-5\mathbf{j}$, $8\mathbf{i}+7\mathbf{j}$
d) (1, 5, 4), (8, 2, 3), (22, −4, 1)

e) (0, 0, 0), (1, 0, 1), (0, 1, 1)
f) (−2, 3, 1), (1, 2, 1), (1, −3, 4).

11. Demuéstrese que $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ colineales son equivalentes a $(\mathbf{P}_2 - \mathbf{P}_1) \times (\mathbf{P}_3 - \mathbf{P}_1) = \mathbf{0}$.

12. Demuéstrese que si $\mathbf{P}_1 \neq \mathbf{P}_2$, entonces $\{\mathbf{P} \mid (\mathbf{P} - \mathbf{P}_1) \times (\mathbf{P}_2 - \mathbf{P}_1) = \mathbf{0}\}$ es la recta que pasa por $\mathbf{P}_1$ y $\mathbf{P}_2$.

13. Sea $\mathscr{L}_1$ la recta que pasa por (1, 0, 1) y (2, 1, 2). Sea $\mathscr{L}_2$ la recta que pasa por el origen y es paralela a (1, 0, 1). Determínese la recta que pasa por el punto (2, 0, −3) ortogonal tanto a $\mathscr{L}_1$ como a $\mathscr{L}_2$.

5. EL TRIPLE PRODUCTO ESCALAR

Dados tres vectores cualesquiera **a**, **b**, y **c** en V_3, entonces como $\mathbf{b} \times \mathbf{c}$ es un vector, podemos formar el producto escalar de **a** con $\mathbf{b} \times \mathbf{c}$. A este producto le llamamos "triple producto escalar".

5.1 Definición. *Dados tres vectores* **a**, **b**, $\mathbf{c} \in V_3$, *el* ***triple producto escalar*** *de* **a**, **b**, **c**, *denotado por* [**abc**], *se define por*

$$[\mathbf{abc}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}).$$

Nótese que expresiones tales como $(\mathbf{a} \cdot \mathbf{b}) \times \mathbf{c}$ no tienen significado alguno, ya que $\mathbf{a} \cdot \mathbf{b}$ es un número real y el producto vectorial es una operación entre pares de vectores de V_3.

El triple producto escalar [**abc**] puede expresarse simplemente en términos de un determinante 3×3:

$$\begin{aligned}
[\mathbf{abc}] &= \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) \\
&= (a_1, a_2, a_3) \cdot (b_2 c_3 - b_3 c_2, b_3 c_1 - b_1 c_3, b_1 c_2 - b_2 c_1) \\
&= a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \\
&= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}
\end{aligned}$$

Mediante el cálculo directo puede demostrarse (problema 1*b*) que

5.2 $$[\mathbf{abc}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}).$$

Como el producto escalar tiene la propiedad conmutativa, la ecuación 5.2 puede reformularse como

5.3 $$[\mathbf{abc}] = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b} = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}.$$

La ecuación 5.2 muestra que el triple producto escalar no cambia por permutaciones cíclicas de los vectores:

$$[\mathbf{abc}] = [\mathbf{bca}] = [\mathbf{cab}]$$

y de la ecuación 5.3 se deduce que al expresar $[\mathbf{abc}]$ podemos colocar el punto y la cruz en cualquiera de las dos posiciones: $[\mathbf{abc}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$.

El triple producto escalar puede usarse para describir la orientación de R^3. Si $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ son tres vectores mutuamente ortogonales y $[\mathbf{abc}] > 0$, entonces decimos que $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ es una terna positivamente orientada. Por ejemplo, los tres vectores unitarios $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ forman una terna positivamente orientada, ya que

$$[\mathbf{ijk}] = \mathbf{i} \cdot (\mathbf{j} \times \mathbf{k}) = \mathbf{i} \cdot \mathbf{i} = 1 > 0.$$

Ya hemos admitido que $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ forman un sistema levógiro. Hemos, pues, convenido en que la orientación levógira la tomaremos como orientación positiva. Luego, si $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ forman una terna positivamente orientada de vectores mutuamente ortogonales, entonces la rotación de $\mathbf{b}$ a $\mathbf{c}$ de un ángulo igual a $\frac{\pi}{2}$ aparece como si fuera contraria a la de las manecillas del reloj cuando se ve desde $\mathbf{a}$.

La noción de terna orientada puede extenderse a cualesquiera tres vectores $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ (no necesariamente ortogonales): $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ constituyen una terna positivamente orientada si $[\mathbf{abc}] > 0$. Consideremos ahora la terna $\mathbf{b} \times \mathbf{c}$, $\mathbf{b}$, $\mathbf{c}$ donde $\mathbf{b}$ y $\mathbf{c}$ no son paralelos. Entonces

$$[(\mathbf{b} \times \mathbf{c})\mathbf{bc}] = (\mathbf{b} \times \mathbf{c}) \cdot (\mathbf{b} \times \mathbf{c}) = |\mathbf{b} \times \mathbf{c}|^2 > 0.$$

Como hemos supuesto que la orientación levógira es la orientación positiva, tenemos que el giro de un ángulo θ de $\mathbf{b}$ a $\mathbf{c}$ donde $0 < \theta < \pi$ parece contrario al movimiento de las manecillas del reloj cuando se ve desde $\mathbf{b} \times \mathbf{c}$ (figura 8).

Para obtener una interpretación geométrica de una terna arbitraria positivamente orientada $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, construimos $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ con el mismo punto inicial $\mathbf{P}_0$ (figura 8) y denominamos $\mathscr{P}$ al plano que pasa por $\mathbf{P}_0$ determinado por $\mathbf{b}$ y $\mathbf{c}$. Como (ecuación 7.4, pág. 33)

$$[\mathbf{abc}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = |\mathbf{b} \times \mathbf{c}| \operatorname{Comp}_{(\mathbf{b} \times \mathbf{c})}\mathbf{a},$$

vemos que la orientación positiva implica que $\operatorname{Comp}_{(\mathbf{b} \times \mathbf{c})} \mathbf{a} > 0$, y, por tanto, que $\mathbf{Proy}_{(\mathbf{b} \times \mathbf{c})}\mathbf{a}$ y $\mathbf{b} \times \mathbf{c}$ apuntan en la misma dirección. Es decir,

si **a**, **b** y **c** forman una terna positivamente orientada, **a** y $\mathbf{b} \times \mathbf{c}$ se encuentran a un mismo lado del plano $\mathscr{P}$.

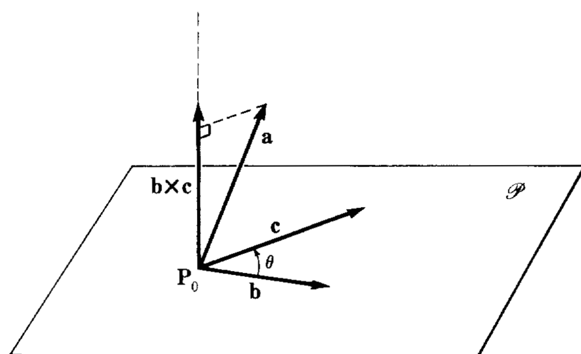


FIGURA 8

Nota. Podíamos haber tomado como orientación positiva la orientación dextrógira. Si hubiésemos hecho tal elección, la única cosa que habría cambiado habrían sido las figuras dibujadas. Por ejemplo, si **b** y **c** no son paralelos, la rotación de **b** a **c** de un ángulo θ donde $0 < \theta < \pi$ habría parecido tener igual dirección que la de las manecillas del reloj cuando es vista desde $\mathbf{b} \times \mathbf{c}$.

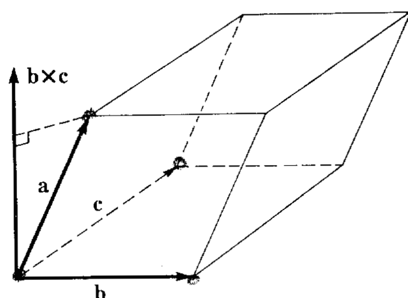


FIGURA 9

Si la terna de vectores **a**, **b**, **c** está positivamente orientada, entonces $[\mathbf{abc}]$ es el volumen del paralelepípedo de lados **a**, **b** y **c** (figura 9). El volumen del paralelepípedo es el área de la base por la altura. La base es un paralelogramo de lados **b** y **c** y, por tanto, su área es $|\mathbf{b} \times \mathbf{c}|$. Ahora bien, la altura es exactamente $\text{Comp}_{(\mathbf{b} \times \mathbf{c})}\mathbf{a}$, luego, por tanto,

$$\text{Volumen} = |\mathbf{b} \times \mathbf{c}| \, \text{Comp}_{(\mathbf{b} \times \mathbf{c})}\mathbf{a} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = [\mathbf{abc}].$$

Si $[\mathbf{abc}] < 0$, entonces $-[\mathbf{abc}]$ es el volumen del paralelepípedo de lados **a**, **b** y **c**.

5.4 Ejemplo. Encuéntrese el volumen del paralelepípedo de lados $\mathbf{a} = (2, 3, -1)$, $\mathbf{b} = (3, -7, 5)$, y $\mathbf{c} = (1, -5, 2)$.

SOLUCIÓN.

$$[\mathbf{abc}] = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \begin{vmatrix} 2 & 3 & -1 \\ 3 & -7 & 5 \\ 1 & -5 & 2 \end{vmatrix} = 27.$$

Luego, el volumen del paralelepípedo es 27.

5.5 Ejemplo. Encuéntrese el volumen del tetraedro de lados **a**, **b** y **c** iguales a los dados en el ejemplo 5.4.

SOLUCIÓN. El volumen de un tetraedro es un tercio del área de la base por la altura. La base es un triángulo con dos lados **b** y **c** y su área es exactamente la mitad del área del paralelogramo de lados **b** y **c**. De donde el área de la base es $\frac{1}{2}|\mathbf{b}\times\mathbf{c}|$ y el volumen del tetraedro es

$$V = \tfrac{1}{3}(\text{área de la base})\,(\text{altura}) = \tfrac{1}{3}\cdot\tfrac{1}{2}|\mathbf{b}\times\mathbf{c}|\ \text{Comp}_{(\mathbf{b}\times\mathbf{c})}\mathbf{a} = \tfrac{1}{6}[\mathbf{abc}].$$

Con **a**, **b** y **c** iguales a los dados en el ejemplo 5.4, tenemos

$$\text{Volumen} = \tfrac{1}{6}\cdot 27 = \tfrac{9}{2}.$$

Problemas

1. Demuéstrese que:

a) $\mathbf{a}\times\mathbf{a} = \mathbf{0}$
b) $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$
c) $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = -\mathbf{b}\cdot(\mathbf{a}\times\mathbf{c})$
d) $\mathbf{a}\cdot(\mathbf{a}\times\mathbf{b}) = 0$.

2. ¿Con qué propiedades de los determinantes se corresponde 1*b*-*d*?

3. Determínense los volúmenes de los paralelepípedos de aristas:

a) $3\mathbf{i}$, $4\mathbf{j}$, $8\mathbf{k}$ *b*) $3\mathbf{i}+4\mathbf{k}$, $\mathbf{i}+\mathbf{k}$, $2\mathbf{j}+4\mathbf{k}$
c) (2, −3, 4), (1,1,1), (1, −4, 7) *d*) (1, 0, 0), (8, 7, 0), (8, −4, 3)
e) (2, 6, −4), (3, 2, 7), (2, 4, 3) *f*) (2, −1, −3), (4, 1, 4), (0, 1, 2).

4. Determínense los volúmenes de los tetraedros de aristas:

a) (2, 2, 4), (1, 5, 2), (1, 0, 1) *b*) (2, 1, 3), (−3, 0, 6), (4, 5, −1)
c) (5, 0, 16), (1, −1, 1), (8, 2, 3) *d*) (1, 5, 4), (1, 1, 0), (1, −3, 4)
e) (2, 6, −4), (1, 1, 1), (1, −4, 3) *f*) (2, 5, −2), (1, 4, 2), (1, 3, 0).

6. INDEPENDENCIA LINEAL DE VECTORES

6.1 Definición. *Un conjunto* $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ *de* k *vectores de* V_n *se dice que es* ***linealmente independiente*** *si*

$$r_1\mathbf{a}_1 + \ldots + r_k\mathbf{a}_k = \mathbf{0} \quad (r_i \in \mathrm{R})$$

implica $r_1 = \ldots = r_k = 0$. *Si* $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ *no es linealmente independiente se dice que es* ***linealmente dependiente***.

Un conjunto de k vectores $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ es, pues, linealmente dependiente si y sólo si hay k números reales $r_1, \ldots, r_k$ no todos iguales a cero, tales que

$$r_1\mathbf{a}_1 + \ldots + r_k\mathbf{a}_k = \mathbf{0}.$$

La expresión $r_1\mathbf{a}_1 + \ldots + r_k\mathbf{a}_k$ donde $r_1, \ldots, r_k \in \mathrm{R}$ se dice que es una *combinación lineal* de los vectores $\mathbf{a}_1, \ldots, \mathbf{a}_k$.

Con frecuencia nos permitiremos ciertas libertades de lenguaje y en lugar de decir que el conjunto $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ es linealmente independiente (o dependiente), diremos que $\mathbf{a}_1, \ldots, \mathbf{a}_k$ son linealmente independientes (o linealmente dependientes).

Si dos vectores $\mathbf{a}$, $\mathbf{b}$ son linealmente dependientes, entonces hay dos números s, t, que no son cero, tales que

$$s\mathbf{a} + t\mathbf{b} = \mathbf{0}.$$

Si $s \neq 0$, entonces $\mathbf{a} = -\frac{t}{s}\mathbf{b}$ mientras que si $t \neq 0$, entonces $\mathbf{b} = -\frac{s}{t}\mathbf{a}$. En cualquier caso, vemos que la dependencia lineal de dos vectores $\mathbf{a}$, $\mathbf{b}$ implica que $\mathbf{a}$, $\mathbf{b}$ son paralelos. Recíprocamente, si $\mathbf{a}$ y $\mathbf{b}$ son paralelos ($\mathbf{a} = r\mathbf{b}$), entonces son linealmente dependientes ($1\mathbf{a} - r\mathbf{b} = \mathbf{0}$). Luego *la dependencia lineal de dos vectores es equivalente a que los dos vectores sean paralelos.*

Si tres vectores $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ de V_3 son linealmente dependientes, entonces hay tres números r, s, t no todos iguales a cero, tales que

$$r\mathbf{a} + s\mathbf{b} + t\mathbf{c} = \mathbf{0}.$$

Si $r \neq 0$, entonces $\mathbf{a} = -\frac{s}{r}\mathbf{b} - \frac{t}{r}\mathbf{c}$ y $\mathbf{a}$ es una combinación lineal de $\mathbf{b}$ y $\mathbf{c}$. Si $\mathbf{b}$ y $\mathbf{c}$ no son paralelos (son linealmente independientes), entonces $\mathbf{b}$ y $\mathbf{c}$ determinan un plano $\mathscr{P}$ que pasa por cualquier punto dado $P_0 \in R^3$ y $\mathbf{a}$ es también paralelo a $\mathscr{P}$. (Decimos que un vector es paralelo a un plano $\mathscr{P}$ si para cualquier punto $\mathbf{P}_0 \in \mathscr{P}$ la recta $\{\mathbf{P}_0 + t\mathbf{a}\} \subset \mathscr{P}$.) Si $\mathbf{b}$ y $\mathbf{c}$ son paralelos (linealmente dependientes), entonces $\mathbf{a}$ es también paralelo a $\mathbf{b}$ y $\mathbf{c}$ y hay muchos planos por cualquier punto $\mathbf{P}_0 \in \mathrm{R}^3$ a los que $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ son paralelos. Análogamente, si $s \neq 0$, entonces $\mathbf{b}$ es una combinación lineal de $\mathbf{a}$ y $\mathbf{c}$, y si $t \neq 0$, entonces $\mathbf{c}$ es una combinación lineal de $\mathbf{a}$ y $\mathbf{b}$. En cualquiera de los casos, hay al menos un plano $\mathscr{P}$ por cualquiera de los puntos $\mathbf{P}_0 \in \mathrm{R}^3$

tal que **a**, **b** y **c** son paralelos a $\mathscr{P}$. Recíprocamente, si tres vectores son paralelos a un mismo plano, puede demostrarse que son linealmente dependientes. De donde *la dependencia lineal de tres vectores es equivalente a que los tres vectores sean paralelos a un mismo plano.*

En la sección 10 se demostrará que cualquier conjunto de más de tres vectores en V_3 es linealmente dependiente.

Si algún subconjunto de un conjunto de k vectores es linealmente dependiente, entonces el conjunto total de k vectores es linealmente dependiente. Supongamos que

6.2 $$r_1\mathbf{a}_1 + \ldots + r_j\mathbf{a}_j = \mathbf{0} \qquad (j \leqslant k)$$

con no todos los r_i iguales a cero y consideremos la ecuación

6.3 $$r_1\mathbf{a}_1 + \ldots + r_j\mathbf{a}_j + r_{j+1}\mathbf{a}_{j+1} + \ldots + r_k\mathbf{a}_k = \mathbf{0}.$$

Deseamos demostrar que es posible escoger coeficientes $r_1, \ldots, r_k$ no todos cero, tales que la ecuación 6.3 se verifica. Podemos escoger $r_1, \ldots, r_j$, no todos cero, tales que la ecuación 6.2 se verifique, y escoger $r_{j+1} = \ldots = r_k = 0$. Entonces tenemos coeficientes $r_1, \ldots, r_k$ para la ecuación 6.3, no todos cero (al menos uno de los números $r_1, \ldots, r_j$ no es cero) y, por tanto, el conjunto $\{\mathbf{a}_1, \ldots, \mathbf{a}_j, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_k\}$ es linealmente dependiente.

Cualquier conjunto de vectores que contiene el vector cero es linealmente dependiente pues podemos escoger coeficientes no todos cero tales que la correspondiente combinación lineal es igual a cero. En particular, podemos tomar todos los coeficientes de los vectores no iguales a cero, iguales a cero y tomar el número uno como coeficiente del vector cero.

Ahora demostraremos que el triple producto escalar nos proporciona un medio conveniente para comprobar la dependencia o independencia lineal de tres vectores en V_3. Como hemos señalado antes, el valor absoluto de [**abc**] es el volumen de un paralelepípedo con aristas **a**, **b**, **c** y es claro, geométricamente, que este volumen es cero si y sólo si **a**, **b** y **c** son paralelos a algún plano. Por otra parte, que tres vectores sean paralelos a un plano es equivalente a la dependencia lineal de los tres vectores.

6.4 Teorema. *Tres vectores* **a**, **b**, $\mathbf{c} \in V_3$ *son linealmente dependientes si y sólo si* $[\mathbf{abc}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = 0$.

PRUEBA. Probamos primero que **a**, **b**, **c** linealmente dependientes implica $[\mathbf{abc}] = 0$. Si **b** y **c** son linealmente dependientes, entonces **a**, **b** y **c** son linealmente dependientes. Pero, entonces, por el teorema 4.10 (pág. 57) $[\mathbf{abc}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{a} \cdot \mathbf{0} = 0$. Si **b** y **c** son linealmente independientes mientras que **a**, **b** y **c** son linealmente dependientes, entonces (problema 4)

$$\mathbf{a} = s\mathbf{b} + t\mathbf{c} \quad \text{para algunos } s, t \in \mathrm{R}.$$

De donde, como tanto **b** como **c** son, ambos, ortogonales a $\mathbf{b}\times\mathbf{c}$,

$$[\mathbf{abc}] = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = s(\mathbf{b}\cdot(\mathbf{b}\times\mathbf{c}))+t(\mathbf{c}\cdot(\mathbf{b}\times\mathbf{c})) = 0.$$

Recíprocamente, supongamos que $[\mathbf{abc}] = 0$. Entonces el vector $\mathbf{b}\times\mathbf{c}$ es ortogonal a **a**, puesto que $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = 0$. Además, $\mathbf{b}\times\mathbf{c}$ es ortogonal a **b**. Por tanto (problema 3, pág. 58), $\mathbf{b}\times\mathbf{c}$ es paralelo a $\mathbf{a}\times\mathbf{b}$. Consideramos dos casos:

Caso 1. Si $\mathbf{a}\times\mathbf{b} = \mathbf{0}$, entonces **a** y **b** son linealmente dependientes y, por tanto, **a**, **b** y **c** son linealmente dependientes.

Caso 2. Si $\mathbf{a}\times\mathbf{b} \neq \mathbf{0}$, entonces para algún $r\in \mathrm{R}$

$$\mathbf{b}\times\mathbf{c} = r(\mathbf{a}\times\mathbf{b}).$$

De donde, reordenando, tenemos

$$\mathbf{b}\times\mathbf{c}+r(\mathbf{b}\times\mathbf{a}) = \mathbf{0},$$

o bien,

$$\mathbf{b}\times(\mathbf{c}+r\mathbf{a}) = \mathbf{0}.$$

Por tanto, **b** y $\mathbf{c}+r\mathbf{a}$ son paralelos. Pero $\mathbf{b} \neq \mathbf{0}$ y, por tanto, para algún $s\in \mathrm{R}$

$$\mathbf{c}+r\mathbf{a} = s\mathbf{b}.$$

Lo que demuestra que **a**, **b** y **c** son linealmente dependientes.

6.5 Ejemplo. ¿Son los vectores $\mathbf{a} = (1, 2, -2)$, $\mathbf{b} = (0, 3, 1)$, y $\mathbf{c} = (-1, 1, 3)$ linealmente dependientes?

SOLUCIÓN. Como

$$[\mathbf{abc}] = \begin{vmatrix} 1 & 2 & -2 \\ 0 & 3 & 1 \\ -1 & 1 & 3 \end{vmatrix}$$

$$= 1\begin{vmatrix} 3 & 1 \\ 1 & 3 \end{vmatrix} - 2\begin{vmatrix} 0 & 1 \\ -1 & 3 \end{vmatrix} - 2\begin{vmatrix} 0 & 3 \\ -1 & 1 \end{vmatrix}$$

$$= 1(9-1)-2(0+1)-2(0+3) = 0,$$

los vectores son linealmente dependientes. En realidad, $\mathbf{a} = \mathbf{b}-\mathbf{c}$.

Problemas

1. Determínese si sí o no son linealmente independientes los siguientes vectores:

a) (2, 5, −1), (3, −7, 0), (0, 29, −3)

b) (3, −7, 5), (6, −5, 2)
c) (1, 1, 1), (1, −1, 1), (1, 1, −1)
d) (12, 52, −9), (2, 6, −1), (1, −5, 2).

2. Demuéstrese que el sistema homogéneo

$$\begin{aligned} a_1 x + b_1 y + c_1 z &= 0 \\ a_2 x + b_2 y + c_2 z &= 0 \\ a_3 x + b_3 y + c_3 z &= 0 \end{aligned}$$

tiene soluciones no triviales [es decir, $(x, y, z) \neq \mathbf{0}$] si y sólo si

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0 .$$

Sugerencia. El sistema de ecuaciones es equivalente a la ecuación vectorial $x\mathbf{a} + y\mathbf{b} + z\mathbf{c} = \mathbf{0}$. Usese el teorema 6.4 y el problema 7*b*, pág. 58.

3. *a*) ¿Para qué valores de λ tiene soluciones no triviales el siguiente sistema de ecuaciones homogéneas?

$$\begin{aligned} (1-\lambda)x + y - z &= 0 \\ 2x - \lambda y - 2z &= 0 \\ x - y - (1+\lambda)z &= 0\,? \end{aligned}$$

b) Determínense las soluciones no triviales para cada uno de los valores de λ que dan lugar a ellas.

4. Demuestrese que si **a**, **b** y **c** son linealmente dependientes mientras que **b** y **c** son linealmente independientes, entonces hay números reales s, t tales que

$$\mathbf{a} = s\mathbf{b} + t\mathbf{c}.$$

5. Demuéstrese que si $\mathbf{a}_1, \ldots, \mathbf{a}_k$ son linealmente independientes, entonces

$$r_1 \mathbf{a}_1 + \ldots + r_k \mathbf{a}_k = s_1 \mathbf{a}_1 + \ldots + s_k \mathbf{a}_k$$

implica $r_1 = s_1, \ldots, r_k = s_k$.

6. Demuéstrese que $\mathbf{a}_1, \ldots, \mathbf{a}_k$ linealmente independientes, implica $\mathbf{a}_1, \ldots, \mathbf{a}_k$ son vectores distintos de cero.

7. Demuéstrese que k vectores son linealmente dependientes si y sólo si uno de los vectores es una combinación lineal de los otros.

***8.** Demuéstrese que un conjunto $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ de vectores no nulos es linealmente dependiente si y sólo si para algún j, $1 \leqslant j \leqslant k-1$, $\mathbf{a}_{j+1}$ es una combinación lineal de $\mathbf{a}_1, \ldots, \mathbf{a}_j$.

***9.** Sea $\mathscr{L}_1$ la recta que pasa por $\mathbf{P}_0$ y es paralela a $\mathbf{a}$. Sea $\mathscr{L}_2$ la recta que pasa por $\mathbf{Q}_0$ y es paralela a $\mathbf{b}$. Sea $\mathbf{c} = \mathbf{Q}_0 - \mathbf{P}_0$. Si $\mathscr{L}_1$ no es paralela a $\mathscr{L}_2$ demuéstrese que:

a) la distancia mínima entre $\mathscr{L}_1$ y $\mathscr{L}_2$ está dada por

$$d = \left| \mathbf{c} \cdot \frac{(\mathbf{a} \times \mathbf{b})}{|\mathbf{a} \times \mathbf{b}|} \right| = \frac{|[\mathbf{abc}]|}{|\mathbf{a} \times \mathbf{b}|}$$

b) las rectas $\mathscr{L}_1$ y $\mathscr{L}_2$ se intersectan si y sólo si $[\mathbf{abc}] = 0$.

10. Establézcanse las siguientes identidades:

a) $(\mathbf{a}+\mathbf{b}) \cdot [(\mathbf{a} \times \mathbf{c}) \times (\mathbf{a}+\mathbf{b})] = 0$
b) $\mathbf{a} \times [\mathbf{a} \times (\mathbf{a} \times \mathbf{b})] = (\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \times \mathbf{a})$
c) $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d}]\mathbf{c} - [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}]\mathbf{d}$
d) $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})$
e) $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{a} \times \mathbf{c}) = [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})]\mathbf{a}$
f) $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$.

11. Demuéstrese que 4.8, pág. 56, es un caso especial del problema 10*d*.

***12.** Sea $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ una terna positivamente orientada de vectores, y sea $\Delta = [\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3]$. Definamos

$$\mathbf{b}_1 = \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\Delta}, \qquad \mathbf{b}_2 = \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\Delta}, \qquad \mathbf{b}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\Delta},$$

y δ_{ij} (llamada delta de Kronecker) por

$$\delta_{ij} = \begin{cases} 1 & \text{para} \quad i = j \\ 0 & \text{para} \quad i \neq j \end{cases}$$

Demuéstrese que:

a) $[\mathbf{b}_1 \mathbf{b}_2 \mathbf{b}_3] = \dfrac{1}{\Delta}$

b) $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ es una terna positivamente orientada de vectores.
c) $\mathbf{a}_i \cdot \mathbf{b}_j = \delta_{ij}$, $i, j = 1, 2, 3$.
d) Los vectores $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_3$ son los únicos vectores con la propiedad *c*.
e) Si $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ es una terna positivamente orientada de vectores unitarios ortogonales dos a dos, entonces $\mathbf{b}_1 = \mathbf{a}_1$, $\mathbf{b}_2 = \mathbf{a}_2$ y $\mathbf{b}_3 = \mathbf{a}_3$.

7. LA ECUACIÓN DEL PLANO

Consideremos el plano (figura 10)

7.1 $$\mathscr{P} = \{\mathbf{P}_0 + u\mathbf{a} + v\mathbf{b} \mid u, v \in \mathrm{R}\}.$$

$\mathscr{P}$ es el plano que pasa por $\mathbf{P}_0$ determinado por el par de vectores no paralelos **a** y **b**. Cualquier vector no nulo ortogonal a ambos **a** y **b** se llama vector *normal* a $\mathscr{P}$. Así pues, $\mathbf{a} \times \mathbf{b}$ es un vector normal (o, simplemente, una normal) a $\mathscr{P}$ y toda otra normal es paralela a $\mathbf{a} \times \mathbf{b}$ (problema 3, pág. 58).

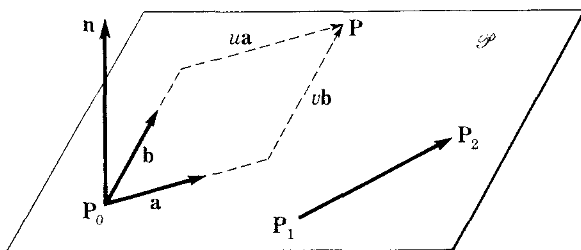


FIGURA 10

7.2 Lema. *Si* **n** *es una normal al plano* $\mathscr{P} = \{\mathbf{P}_0 + u\mathbf{a} + v\mathbf{b} \mid u, v \in \mathrm{R}\}$ *y* $\mathbf{P}_1, \mathbf{P}_2 \in \mathscr{P}$ *entonces* **n** *es ortogonal a* $\mathbf{P}_2 - \mathbf{P}_1$.

Prueba. $\mathbf{P}_1, \mathbf{P}_2 \in \mathscr{P}$ implica

$$\mathbf{P}_1 = \mathbf{P}_0 + u_1\mathbf{a} + v_1\mathbf{b} \quad \text{y} \quad \mathbf{P}_2 = \mathbf{P}_0 + u_2\mathbf{a} + v_2\mathbf{b}$$

para algunas u_1, v_1, u_2, v_2 de R. Por tanto

$$\mathbf{P}_2 - \mathbf{P}_1 = (u_2 - u_1)\mathbf{a} + (v_2 - v_1)\mathbf{b}.$$

Como **n** es ortogonal tanto a **a** como a **b**, es claro que

$$\mathbf{n} \cdot (\mathbf{P}_2 - \mathbf{P}_1) = 0.$$

Y esto completa la prueba.

7.3 Lema. *Si* **n** *es una normal al plano* $\mathscr{P} = \{\mathbf{P}_0 + u\mathbf{a} + v\mathbf{b} \mid u, v \in \mathrm{R}\}$ *y* $\mathbf{P} - \mathbf{P}_0$ *es ortogonal a* **n**, *entonces* $\mathbf{P} \in \mathscr{P}$.

Prueba. Como $\mathbf{n} = r(\mathbf{a} \times \mathbf{b})$ y $r \neq 0$, $\mathbf{P} - \mathbf{P}_0$ ortogonal a **n** implica

$$(\mathbf{P} - \mathbf{P}_0) \cdot (\mathbf{a} \times \mathbf{b}) = 0.$$

De acuerdo con el teorema 6.4 (pág. 64), esto implica que $\mathbf{P} - \mathbf{P}_0$, **a** y **b** son linealmente dependientes. Como **a** y **b** son linealmente independientes, concluimos que existen $u, v \in \mathrm{R}$ tales que

$$\mathbf{P} - \mathbf{P}_0 = u\mathbf{a} + v\mathbf{b}.$$

Por tanto, $\mathbf{P} = \mathbf{P}_0 + u\mathbf{a} + v\mathbf{b}$ y $\mathbf{P} \in \mathscr{P}$.

De los lemas 7.2 y 7.3 se deduce

7.4 Teorema. *Si* **n** *es una normal al plano*

$$\mathscr{P} = \{\mathbf{P}_0 + u\mathbf{a} + v\mathbf{b} \mid u, v \in \mathrm{R}\}$$

entonces

$$\mathscr{P} = \{\mathbf{P} \mid \mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = 0\}$$

y $\mathscr{P}$ *es el único plano que pasa por* $\mathbf{P}_0$ *con normal* **n**.

PRUEBA. Sea $\mathscr{S} = \{\mathbf{P} \mid \mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = 0\}$. Deseamos demostrar que $\mathscr{S} = \mathscr{P}$. Según el lema 7.2, si $\mathbf{P} \in \mathscr{P}$, entonces $\mathbf{P} - \mathbf{P}_0$ es ortogonal a **n** y $\mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = 0$. De donde $\mathbf{P} \in \mathscr{P}$ implica $\mathbf{P} \in \mathscr{S}$ de modo que $\mathscr{P} \subset \mathscr{S}$. Recíprocamente, si $\mathbf{P} \in \mathscr{S}$ entonces $\mathbf{P} - \mathbf{P}_0$ es ortogonal a **n** y según el lema 7.3 $\mathbf{P} \in \mathscr{P}$. Por tanto $\mathscr{S} \subset \mathscr{P}$. Luego $\mathscr{S} = \mathscr{P}$.

Para demostrar que $\mathscr{P}$ es el único plano que pasa por $\mathbf{P}_0$ con normal **n**, supongamos que

$$\mathscr{P}' = \{\mathbf{P}_0' + s\mathbf{c} + t\mathbf{d} \mid s, t \in \mathrm{R}\}$$

es otro plano que pase por $\mathbf{P}_0$ y tenga también a **n** como normal. Entonces

$$\mathscr{P}' = \{\mathbf{P} \mid \mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0') = 0\}.$$

Como $\mathbf{P}_0 \in \mathscr{P}'$, $\mathbf{n} \cdot (\mathbf{P}_0 - \mathbf{P}_0') = 0$ o, lo que es lo mismo, $\mathbf{n} \cdot \mathbf{P}_0 = \mathbf{n} \cdot \mathbf{P}_0'$. De donde $\mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = \mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0')$ para todo $\mathbf{P} \in \mathrm{R}^3$ y en particular

$$\mathscr{P}' = \{\mathbf{P} \mid \mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0') = 0\} = \{\mathbf{P} \mid \mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = 0\} = \mathscr{P}.$$

Y esto completa la prueba.

La ecuación

7.5 $$\mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = 0$$

se llama *ecuación vectorial del plano* $\mathscr{P}$.

Hemos demostrado que si $\mathscr{P}$ es un plano que pasa por $\mathbf{P}_0$ y tiene **n** como normal, entonces $\mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = 0$ es una ecuación vectorial de $\mathscr{P}$. Ahora demostraremos que, recíprocamente, toda ecuación vectorial $\mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = 0$ $(\mathbf{n} \neq \mathbf{0})$ es la ecuación vectorial de un plano que pasa por $\mathbf{P}_0$.

7.6 Teorema. *Para todo vector distinto de cero* **n** *y todo punto* $\mathbf{P}_0$, $\mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = 0$, *es una ecuación vectorial de un plano que pasa por* $\mathbf{P}_0$ *y tiene* **n** *como normal.*

PRUEBA. Deseamos demostrar que

$$\mathscr{S} = \{\mathbf{P} \mid \mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = 0\}$$

es un plano que pasa por $\mathbf{P}_0$ y tiene **n** como normal. Necesitamos demostrar

tan solo que existen vectores linealmente independientes **a** y **b** ambos ortogonales a **n**. Entonces, **n** es una normal al plano

$$\mathscr{P} = \{\mathbf{P}_0 + u\mathbf{a} + v\mathbf{b} \mid u, v \in \mathrm{R}\}$$

y, por el teorema 7.4, $\mathscr{S} = \mathscr{P}$.

Sea $\mathbf{n} = (n_1, n_2, n_3)$. Como $\mathbf{n} \neq \mathbf{0}$, al menos uno de sus componentes es distinto de cero. Supongamos que $n_1 \neq 0$. Entonces

$$\mathbf{a} = (-n_2, n_1, 0)$$

es un vector distinto de cero ortogonal a **n**. Como **a** y **n** son vectores ortogonales distintos de cero, son linealmente independientes (problema 10, pág. 29). Entonces $\mathbf{b} = \mathbf{n} \times \mathbf{a}$ es un vector distinto de cero ortogonal a **n** y a **a**. Luego **a** y **b** son vectores linealmente independientes cada uno de los cuales es ortogonal a **n**. Lo que completa la prueba.

7.7 Corolario. *Toda ecuación vectorial* $\mathbf{n} \cdot \mathbf{P} = d(\mathbf{n} \neq \mathbf{0})$ *es una ecuación de un plano que tiene* **n** *como normal.*

PRUEBA. Si la ecuación $\mathbf{n} \cdot \mathbf{P} = d$ tiene una solución $\mathbf{P}_0$, entonces $\mathbf{n} \cdot \mathbf{P}_0 = d$. De donde $\mathbf{n} \cdot \mathbf{P} = d = \mathbf{n} \cdot \mathbf{P}_0$ es equivalente a $\mathbf{n} \cdot (\mathbf{P} - \mathbf{P}_0) = 0$, y sabemos que esta es una ecuación de un plano. Queda, pues, por probar que la ecuación $\mathbf{n} \cdot \mathbf{P} = d$ tiene una solución. Sea $\mathbf{n} = (a, b, c)$. Como $\mathbf{n} \neq \mathbf{0}$, al menos uno de sus componentes es distinto de cero. Deseamos, pues, encontrar un punto $\mathbf{P} = (x, y, z)$ que satisfaga a

$$ax + by + cz = d.$$

Claramente, como al menos uno de los números a, b, c es distinto de cero, la ecuación tiene soluciones [por ejemplo, $(d/a, 0, 0)$ si $a \neq 0$; $(0, d/b, 0)$ si $b \neq 0$; $(0, 0, d/c)$ si $c \neq 0$]. Y esto completa la prueba.

Como en la prueba anterior, sea $\mathbf{n} = (a, b, c) \neq \mathbf{0}$ y $\mathbf{P} = (x, y, z)$. Entonces

$$\mathbf{n} \cdot \mathbf{P} = ax + by + cz,$$

y escrita en términos de componentes la ecuación vectorial $\mathbf{n} \cdot \mathbf{P} = d$ toma la forma

7.8 $$ax + by + cz = d.$$

Así pues, el conjunto $\mathscr{P} = \{(x, y, z) \mid ax + by + cz = d\}$ de todas las soluciones a 7.8 es un plano con normal $\mathbf{n} = (a, b, c)$; 7.8 se llama *ecuación del plano* $\mathscr{P}$. Una ecuación del tipo 7.8, donde $a^2 + b^2 + c^2 \neq 0$, se llama *ecuación lineal* en x, y, z. *Luego toda ecuación lineal en* x, y, z *es la ecuación de un plano.*

7.9 Ejemplo. Identifíquense cada uno de los siguientes planos:

a) $3(x-5)-2(y+4)+4(z-2) = 0$

b) $2x+3y = 2$

c) $x-2y+z = 0$.

Solución. Un plano queda inequívocamente determinado por una normal $\mathbf{n}$ y un punto $\mathbf{P}_0$ del plano. De donde $\mathbf{n}$ y $\mathbf{P}_0$ identifican un plano.

a) $\mathbf{n} = (3, -2, 4)$, $\mathbf{P}_0 = (5, -4, 2)$.

b) (En dos dimensiones esto es una recta, pero se entiende que aquí estamos discutiendo la geometría del espacio euclidiano tridimensional.)

$$\mathbf{n} = (2, 3, 0) \quad \text{y} \quad \mathbf{P}_0 = (1, 0, 0).$$

c) $\mathbf{n} = \mathbf{i}-2\mathbf{j}+\mathbf{k}$, y el plano pasa por el origen.

7.10 Ejemplo. Determínese la recta que pasa por el punto $(1, -5, 6)$ paralela a la normal al plano que contiene a los puntos $(0, 1, 2)$, $(3, 2, 6)$ y $(-2, 0, 5)$.

Solución. Sea $\mathbf{a} = (3, 2, 6)-(0, 1, 2) = (3, 1, 4)$ y $\mathbf{b} = (-2, 0, 5)-(0, 1, 2) = (-2, -1, 3)$. $\mathbf{n} = \mathbf{a}\times\mathbf{b}$ es una normal a $\mathscr{P}$. Luego

$$\mathbf{n} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 1 & 4 \\ -2 & -1 & 3 \end{vmatrix} = 7\mathbf{i}-17\mathbf{j}-\mathbf{k} = (7, -17, -1).$$

De donde $\mathscr{L} = \{(1, -5, 6)+t(7, -17, -1)\} = \{(1+7t, -5-17t, 6-t)\}$ es la recta.

7.11 Teorema. *Tres puntos no colineales determinan un plano único.*

Prueba. Sean $\mathbf{P}_0$, $\mathbf{P}_1$ y $\mathbf{P}_2$ tres puntos no colineales. Entonces $\mathbf{P}_1-\mathbf{P}_0$ y $\mathbf{P}_2-\mathbf{P}_0$ son linealmente independientes (problema 6, pág. 48). Luego

$$\mathscr{P} = \{\mathbf{P}_0+u(\mathbf{P}_1-\mathbf{P}_0)+v(\mathbf{P}_2-\mathbf{P}_0)\}$$

es un plano que pasa por los puntos $\mathbf{P}_0$, $\mathbf{P}_1$ y $\mathbf{P}_2$. Ahora bien,

$$\mathbf{n} = (\mathbf{P}_1-\mathbf{P}_0)\times(\mathbf{P}_2-\mathbf{P}_0)$$

es una normal a todo plano que pase por $\mathbf{P}_0$, $\mathbf{P}_1$ y $\mathbf{P}_2$ (lema 7.2 y problema 3, pág. 58). Luego de acuerdo con el teorema 7.4, $\mathscr{P}$ es el único plano que pasa por los puntos no colineales $\mathbf{P}_0$, $\mathbf{P}_1$ y $\mathbf{P}_2$.

Problemas

1. Determínese una normal a cada uno de los siguientes planos.

a) El plano cuyas ecuaciones paramétricas son

$$x = 2-3t+s$$
$$y = 8t+7s$$
$$z = -4+3s.$$

b) El plano

$$\mathscr{P} = \{(6, t, s-t) \mid s, t \in \mathrm{R}\}.$$

c) El plano

$$\mathscr{P} = \{(6-u+3v, 8+2u+3v, -1+v) \mid u, v \in \mathrm{R}\}.$$

d) El plano que pasa por los puntos (1, 0, 0), (0, 1, 0), (0, 0, 1).
e) El plano que pasa por los puntos (2, −1, 3), (11, −13, 6), (5, 5, 5).

2. Determínese una ecuación para cada uno de los siguientes planos:

a) El plano que pasa por el origen con normal (1, 1, 1).
b) El plano que pasa por el punto (0, 0, a) con normal paralela al eje Z.
c) El plano que pasa por el punto (1, −4, 3) con normal paralela a la recta que pasa por (2, −1, 3) y (4, 8, 0).
d) El plano que contiene la recta $\mathscr{L} = \{1, 2+3t, 2+t\}$ y el punto (2, −3, 8).
e) El plano que pasa por el punto medio del segmento rectilíneo que une $\mathbf{P}_1$ y $\mathbf{P}_2$ con normal paralela a dicho segmento.

3. Proporciónese una ecuación para cada uno de los planos de problema 1.

4. Identifíquese el plano cuya ecuación es:

a) $3x-5y+z = 0$
b) $3(x-2)+5(y+3)-8z = 0$
c) $z = 6$
d) $2x+3y-8z = 13$
e) $5x-7y+12z = -8$.

***5.** Demuéstrese que el lugar geométrico de todos los puntos equidistantes de un par de puntos distintos es un plano.

6. Demuéstrese que $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ son coplanares si y sólo si $\mathbf{a}_1 = \mathbf{P}_1-\mathbf{P}_0$, $\mathbf{a}_2 = \mathbf{P}_2-\mathbf{P}_0$ y $\mathbf{a}_3 = \mathbf{P}_3-\mathbf{P}_0$ son linealmente dependientes.

Sugerencia. Demuéstrese que $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ coplanares es equivalente a $\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3) = 0$.

8. INTERSECCIÓN DE PLANOS

En esta sección discutiremos el problema de la determinación de la intersección de dos planos. El carácter de esta intersección depende de que los planos sean o no paralelos.

8.1 Definición. *Se dice que dos planos son* ***paralelos*** *si sus normales son paralelas.*

Nota. Si dos planos paralelos $\mathscr{P}_1$ y $\mathscr{P}_2$ tienen normales $\mathbf{n}_1$ y $\mathbf{n}_2$ respectivamente, entonces $\mathbf{n}_1$ y $\mathbf{n}_2$ son vectores paralelos no nulos. Ahora bien, como cualquier vector no nulo paralelo a $\mathbf{n}_1$ es normal a $\mathscr{P}_1$, $\mathbf{n}_2$ es también una normal a $\mathscr{P}_1$. Análogamente, $\mathbf{n}_1$ es una normal a $\mathscr{P}_2$. Es decir, toda normal a uno de los planos de un par de planos paralelos es una *normal común* a los dos planos.

Probaremos ahora que planos paralelos o coinciden o no tienen puntos en común, y que planos no paralelos se intersectan en una recta.

8.2 Teorema. *Si $\mathscr{P}_1$ y $\mathscr{P}_2$ son planos paralelos, entonces $\mathscr{P}_1 = \mathscr{P}_2$ o bien $\mathscr{P}_1 \cap \mathscr{P}_2 = \varnothing$. Si $\mathscr{P}_1$ y $\mathscr{P}_2$ no son paralelos, entonces $\mathscr{P}_1 \cap \mathscr{P}_2$ es una recta.*

PRUEBA. Sean $\mathscr{P}_1 = \{\mathbf{P}_1 + u\mathbf{a} + v\mathbf{b} \mid u, v \in R\}$ y $\mathscr{P}_2 = \{\mathbf{P} \mid (\mathbf{P} - \mathbf{P}_2) \cdot \mathbf{n} = 0\}$ Un punto $\mathbf{P} = \mathbf{P}_1 + u\mathbf{a} + v\mathbf{b}$ en $\mathscr{P}_1$ se encuentra también en $\mathscr{P}_2$ si y sólo si

$$(\mathbf{P}_1 + u\mathbf{a} + v\mathbf{b} - \mathbf{P}_2) \cdot \mathbf{n} = 0$$

o bien

8.3 $$(\mathbf{a} \cdot \mathbf{n})u + (\mathbf{b} \cdot \mathbf{n})v = (\mathbf{P}_2 - \mathbf{P}_1) \cdot \mathbf{n}.$$

Si los planos $\mathscr{P}_1$ y $\mathscr{P}_2$ son paralelos, entonces $\mathbf{a} \cdot \mathbf{n} = 0$ y $\mathbf{b} \cdot \mathbf{n} = 0$. En este caso o bien ningún par de números $u, v \in R$ satisface la ecuación 8.3, o bien cualesquier números $u, v \in R$ la satisfacen, según que $(\mathbf{P}_2 - \mathbf{P}_1) \cdot \mathbf{n}$ sea distinto de cero o sea igual a cero. Si $(\mathbf{P}_2 - \mathbf{P}_1) \cdot \mathbf{n} \neq 0$, entonces no hay números u, v que satisfagan la ecuación 8.3 y $\mathscr{P}_1 \cap \mathscr{P}_2 = \varnothing$. Si $(\mathbf{P}_2 - \mathbf{P}_1) \cdot \mathbf{n} = 0$, entonces cualesquier números u, v satisfacen la ecuación 8.3 y $\mathscr{P}_1 \subset \mathscr{P}_2$. De lo que se sigue que $\mathscr{P}_1 = \mathscr{P}_2$ ya que tres puntos no colineales determinan un plano (teorema 7.11).

Si $\mathscr{P}_1$ y $\mathscr{P}_2$ son no paralelos, entonces $\mathbf{a} \cdot \mathbf{n} \neq 0$ o $\mathbf{b} \cdot \mathbf{n} \neq 0$. Supongamos que $\mathbf{a} \cdot \mathbf{n} \neq 0$. Entonces la ecuación 8.3 puede resolverse para u en términos de v y para cada $v \in R$ tenemos

8.4 $$u = \frac{(\mathbf{P}_2 - \mathbf{P}_1) \cdot \mathbf{n} - (\mathbf{b} \cdot \mathbf{n})v}{\mathbf{a} \cdot \mathbf{n}}.$$

Así pues, un punto $\mathbf{P} = \mathbf{P}_1 + u\mathbf{a} + v\mathbf{b}$ está en $\mathscr{P}_1 \cap \mathscr{P}_2$ si y sólo si u está determinada por 8.4. Es decir,

$$\mathscr{P}_1 \cap \mathscr{P}_2 = \left\{ \mathbf{P}_1 + \frac{(\mathbf{P}_2 - \mathbf{P}_1)\cdot\mathbf{n}}{\mathbf{a}\cdot\mathbf{n}}\mathbf{a} + \left(\mathbf{b} - \frac{\mathbf{b}\cdot\mathbf{n}}{\mathbf{a}\cdot\mathbf{n}}\mathbf{a}\right)v \mid v \in \mathrm{R} \right\}$$

y este conjunto es una recta, ya que $\mathbf{a}$ y $\mathbf{b}$ no paralelos implica $\mathbf{b} - \dfrac{\mathbf{b}\cdot\mathbf{n}}{\mathbf{a}\cdot\mathbf{n}}\mathbf{a} \neq \mathbf{0}$.

8.5 Ejemplo. Encuéntrese la intersección del plano

$$\mathscr{P}_1 = \{(1, 1, 1) + u(2, -1, 3) + v(-1, 0, 2) \mid u, v \in \mathrm{R}\}$$

y el plano $\mathscr{P}_2$ cuya ecuación es

$$2x + 3y - z = 7.$$

SOLUCIÓN. Un punto

$$\mathbf{P} = (1, 1, 1) + u(2, -1, 3) + v(-1, 0, 2) = (1 + 2u - v, 1 - u, 1 + 3u + 2v)$$

en $\mathscr{P}_1$ se encuentra también en $\mathscr{P}_2$ si y sólo si

$$\mathbf{P}\cdot\mathbf{n} = (1 + 2u - v, 1 - u, 1 + 3u + 2v)\cdot(2, 3, -1) = 7$$

o bien

$$4 - 2u - 4v = 7.$$

Luego

$$u = -2v - \tfrac{3}{2}$$

y

$$\begin{aligned}\mathbf{P} &= (1, 1, 1) - \tfrac{3}{2}(2, -1, 3) - 2v(2, -1, 3) + v(-1, 0, 2)\\ &= (-2, \tfrac{5}{2}, -\tfrac{7}{2}) + v(-5, 2, -4).\end{aligned}$$

La intersección de $\mathscr{P}_1$ y $\mathscr{P}_2$ es la recta

$$\mathscr{P}_1 \cap \mathscr{P}_2 = \{(-2, \tfrac{5}{2}, -\tfrac{7}{2}) + v(-5, 2, -4) \mid v \in \mathrm{R}\}.$$

Si los dos planos están dados en forma de ecuación, entonces podemos encontrar mejor la intersección expresando dos de las incógnitas en términos de la tercera. La tercera incógnita pasa entonces a desempeñar el papel de parámetro en la recta de intersección.

Nota. Si P y Q son dos proposiciones, entonces "P si y sólo si Q" significa que P implica Q y Q implica P. Es decir, "P si y sólo si Q" significa que la proposición P es equivalente a la proposición Q. Es frecuente utilizar una flecha, $\Rightarrow$, para denotar "implica"; una flecha de dos cabezas, $\Leftrightarrow$, denota "si y sólo si".

8.6 Ejemplo. Encuéntrense los puntos de intersección de los dos planos $4x + 3y + z = 0$ y $x + y - z = 15$.

SOLUCIÓN.

$$\begin{cases} 4x+3y+z = 0 \\ x + y - z = 15 \end{cases}$$

$$\Leftrightarrow \begin{cases} x = -y+z+15 \\ -4y+4z+60+3y+z = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x = -y+z+15 \\ y = 5z+60 \end{cases}$$

$$\Leftrightarrow \begin{cases} x = -5z-60+z+15 = -4z-45 \\ y = 5z+60. \end{cases}$$

Por tanto, los dos planos se intersectan a lo largo de la recta cuyas ecuaciones paramétricas son

$$\begin{aligned} x &= -4t-45 \\ y &= 5t+60 \\ z &= t. \end{aligned}$$

COMPROBACIÓN.

$$\begin{aligned} 4(-4t-45)+3(5t+60)+t &= 0 \\ (-4t-45)+(5t+60)-t &= 15. \end{aligned}$$

8.7 Definición. *Un **ángulo entre dos planos** es un ángulo entre sus normales.*

8.8 Ejemplo. Encuéntrese un ángulo entre los dos planos del ejemplo 8.6.

SOLUCIÓN. Los planos

$$\begin{aligned} 4x+3y+z &= 0 \\ x+\ \ y-z &= 15 \end{aligned}$$

tienen normales $\mathbf{n}_1 = (4, 3, 1)$ y $\mathbf{n}_2 = (1, 1, -1)$ respectivamente. Por tanto, si θ es un ángulo entre los dos planos, entonces

$$\cos\theta = \frac{\mathbf{n}_1 \cdot \mathbf{n}_2}{|\mathbf{n}_1|\,|\mathbf{n}_2|} = \frac{4\cdot 1+3\cdot 1+1\cdot(-1)}{\sqrt{4^2+3^2+1^2}\,\sqrt{1^2+1^2+1^2}} = \frac{6}{\sqrt{26}\,\sqrt{3}} = 0.6794$$

y $\theta = 47°12'$ o $312°48'$.

Problemas

1. Determínese la intersección de los planos:

a) $7x+2y-8z = 0$
$y+z = 0$

b) $3x-2y+5z = 2$
$4x+5y+z = -6$

c) $12x-5y+7z = 13$
$9x+y-3z = 5$

d) $9x+12y+3z = -7$
$12x+16y+4z = 9$

e) $x+y+z = 1$
$x-y+z = 0$
$x+y-z = 1$

f) $x+y+z = 0$
$x-y+z = 0$
$x+y-z = 0$

g) $3x+2y+z = 2$
$x-2y+3z = 0$
$4x+y+8z = 12$

h) $x-y+z = 1$
$x+y+z = 0$
$x-9y+z = 2.$

2. Determínense los ángulos entre los planos:

a) del problema 1 *a*;
b) del problema 1 *b*;
c) del problema 1 *c*;
d) del problema 1 *d*.

3. Determínese el ángulo entre el plano que pasa por los puntos (1, 0, 0), (0, 1, 0), (0, 0, 1) y el plano cuya ecuación es $3x-5y+z = 8$.

9. INTERSECCIÓN DE UNA RECTA Y UN PLANO

Para discutir la intersección de una recta y un plano, introducimos primero el concepto de recta paralela a un plano.

9.1 Definición. *Una recta $\mathscr{L}$ es paralela a un plano $\mathscr{P}$ si $\mathscr{L}$ es ortogonal a una normal a $\mathscr{P}$.*

9.2 Teorema. *Si la recta $\mathscr{L}$ es paralela al plano $\mathscr{P}$, entonces $\mathscr{L} \subset \mathscr{P}$ o $\mathscr{L} \cap \mathscr{P} = \varnothing$. Si $\mathscr{L}$ no es paralela a $\mathscr{P}$, entonces $\mathscr{L} \cap \mathscr{P}$ contiene un punto.*

PRUEBA. Sean $\mathscr{L} = \{\mathbf{P}_1 + t\mathbf{a} \mid t \in \mathrm{R}\}$ y $\mathscr{P} = \{\mathbf{P} \mid \mathbf{P} \cdot \mathbf{n} = d\}$. Un punto $\mathbf{P} = \mathbf{P}_1 + t\mathbf{a}$ en $\mathscr{L}$ también se encuentra en $\mathscr{P}$ si y sólo si

$$(\mathbf{P}_1 + t\mathbf{a}) \cdot \mathbf{n} = d$$

o bien

9.3 $$(\mathbf{a} \cdot \mathbf{n})t = d - \mathbf{P}_1 \cdot \mathbf{n}.$$

Si $\mathscr{L}$ es paralela a $\mathscr{P}$, entonces **a** y **n** son ortogonales y $\mathbf{a} \cdot \mathbf{n} = 0$. En este caso y según que $d - \mathbf{P}_1 \cdot \mathbf{n}$ sea distinto de cero o cero, respectivamente, o ningún número t o cualquier número t satisfacen la ecuación 9.3. Si $d - \mathbf{P}_1 \cdot \mathbf{n} \neq 0$, entonces ningún $t \in \mathrm{R}$ satisface la ecuación 9.3 y $\mathscr{L} \cap \mathscr{P} = \varnothing$. Si $d - \mathbf{P}_1 \cdot \mathbf{n} = 0$, entonces toda $t \in \mathrm{R}$ satisface a la ecuación 9.3 y $\mathscr{L} \subset \mathscr{P}$.

Si $\mathscr{L}$ no es paralela a $\mathscr{P}$, entonces **a** y **n** no son ortogonales y $\mathbf{a} \cdot \mathbf{n} \neq 0$. En este caso la ecuación 9.3 tiene la solución única

$$t = \frac{d - \mathbf{P}_1 \cdot \mathbf{n}}{\mathbf{a} \cdot \mathbf{n}}$$

y el punto de intersección es

$$\mathbf{P}_1 + \frac{d - \mathbf{P}_1 \cdot \mathbf{n}}{\mathbf{a} \cdot \mathbf{n}} \mathbf{a}.$$

9.4 Ejemplo. Encuéntrese la intersección de la recta

$$\mathscr{L} = \{(1, 1, 1) + t(2, -1, 3) \mid t \in \mathrm{R}\}$$

y el plano $\mathscr{P}$ cuya ecuación es

$$2x + 3y - z = 7.$$

SOLUCIÓN. Un punto $\mathbf{P} = (1, 1, 1) + t(2, -1, 3)$ en $\mathscr{L}$ también se encuentra en $\mathscr{P}$ si y sólo si

$$\mathbf{P} \cdot \mathbf{n} = (1 + 2t, 1 - t, 1 + 3t) \cdot (2, 3, -1) = 7$$

o bien

$$4 - 2t = 7.$$

Así pues, $t = -\frac{3}{2}$ y el punto de intersección es

$$(1, 1, 1) - \tfrac{3}{2}(2, -1, 3) = (-2, \tfrac{5}{2}, -\tfrac{7}{2}).$$

9.5 Definición. *La distancia de un punto* **Q** *a un plano* $\mathscr{P}$ *es la distancia de* **Q** *al punto de intersección con* $\mathscr{P}$ *de la recta que pasa por* **Q** *y es normal a* $\mathscr{P}$.

9.6 Ejemplo. Encuéntrese la distancia de un punto **Q** a un plano $\mathscr{P}$.

SOLUCIÓN. Sea $\mathbf{n} \cdot \mathbf{P} = d$ una ecuación de $\mathscr{P}$. La recta $\mathscr{L}$ que pasa por **Q** normal a $\mathscr{P}$ tiene la ecuación

$$\mathbf{P} = \mathbf{Q} + t\mathbf{n}.$$

De acuerdo con el teorema 9.2, el punto de intersección de $\mathscr{P}$ y $\mathscr{L}$ es

$$\mathbf{P}_1 = \mathbf{Q} + t_1 \mathbf{n} = \mathbf{Q} + \frac{d - \mathbf{n} \cdot \mathbf{Q}}{\mathbf{n} \cdot \mathbf{n}} \mathbf{n}$$

y la distancia es

9.7 $$d(\mathbf{Q}, \mathscr{P}) = |\mathbf{Q} - \mathbf{P}_1| = \left| \frac{d - \mathbf{n} \cdot \mathbf{Q}}{\mathbf{n} \cdot \mathbf{n}} \mathbf{n} \right| = \frac{|d - \mathbf{n} \cdot \mathbf{Q}|}{|\mathbf{n}|}$$

Si $\mathbf{P}_0 \in \mathscr{P}$, entonces $d = \mathbf{n} \cdot \mathbf{P}_0$ y podemos expresar la distancia en la forma

$$d(\mathbf{Q}, \mathscr{P}) = \frac{|\mathbf{n} \cdot (\mathbf{Q} - \mathbf{P}_0)|}{|\mathbf{n}|} = |\mathbf{Proy}_{\mathbf{n}} (\mathbf{Q} - \mathbf{P}_0)|.$$

9.8 Ejemplo. Encuéntrese la distancia del punto $\mathbf{Q} = (1, -2, 4)$ al plano $\mathscr{P} : 2x - y + 3z = 10$.

SOLUCIÓN 1. $\mathbf{n} = (2, -1, 3)$ es una normal a $\mathscr{P}$. Usando la ecuación 9.7, tenemos

$$d(\mathbf{Q}, \mathscr{P}) = \frac{|10 - (2, -1, 3) \cdot (1, -2, 4)|}{\sqrt{4 + 1 + 9}} = \frac{6}{\sqrt{14}}.$$

SOLUCIÓN 2. $\mathbf{n} = (2, -1, 3)$ es una normal a $\mathscr{P}$. Sea $\mathbf{P}_0$ un punto del plano y sea $\mathbf{a} = \mathbf{Q} - \mathbf{P}_0$. Entonces

$$d(\mathbf{Q}, \mathscr{P}) = |\mathbf{Proy}_{\mathbf{n}}\, \mathbf{a}| = \left| \frac{\mathbf{a} \cdot \mathbf{n}}{|\mathbf{n}|} \right| = \frac{|\mathbf{Q} \cdot \mathbf{n} - \mathbf{P}_0 \cdot \mathbf{n}|}{|\mathbf{n}|}$$

$$= \frac{|(1, -2, 4) \cdot (2, -1, 3) - 10|}{\sqrt{14}} = \frac{6}{\sqrt{14}}.$$

Problemas

1. Determínese en cada caso la intersección de $\mathscr{L}$ y $\mathscr{P}$ y dígase si $\mathscr{L}$ es o no paralela a $\mathscr{P}$.

a) $\mathscr{L} = \{(2, 1, 4) + t(1, 1, 1)\}$, $\mathscr{P} = \{(2, 0, 4) + u(1, 7, 3) + v(-3, 8, 0)\}$
b) $\mathscr{L} = \{(1, -1, 4) + t(2, -1, 3)\}$, $\mathscr{P} = \{(6, u, v - u)\}$
c) $\mathscr{L} = \{(3, 8, -1) + t(1, 7, 1)\}$,
$\mathscr{P} = \{(6 - u + 3v, 8 + 2u + 3v, -1 + v)\}$
d) $\mathscr{L} = \{(3, -2, 7) + t(2, -1, 3)\}$, $\mathscr{P}$ es el plano que pasa por los puntos $(2, -1, 3)$, $(5, -5, 4)$, $(5, 5, 8)$
e) $\mathscr{L} = \{(3, 2, 3) + t(-2, -2, -2)\}$, $\mathscr{P}$ es el plano que pasa por el origen con normal $(1, 1, 1)$.

2. Encuéntrese en cada caso la recta $\mathscr{L}$ que pasa por el punto $\mathbf{Q}$ y es ortogonal al plano $\mathscr{P}$.

a) $\mathbf{Q} = (1, 2, 3)$, $\mathscr{P} = \{(2, 1, -1) + u(1, 1, 1) + v(-1, 1, 0)\}$
b) $\mathbf{Q} = (2, 1, -1)$, $\mathscr{P} = \{(2, 1, 3) + u(5, 2, -1) + v(4, 0, 1)\}$
c) $\mathbf{Q} = (0, 2, -2)$, $\mathscr{P}$ que pasa por $(2, 1, -1)$, $(3, 1, 0)$, $(4, -6, 2)$
d) $\mathbf{Q} = (1, -1, 4)$, $\mathscr{P} : 2x + y + z = 5$.

3. Encuéntrese la distancia del punto $\mathbf{Q}$ al plano $\mathscr{P}$ en cada uno de los casos del problema 2.

4. Encuéntrese la intersección de la recta

$$\mathscr{L} = \{(3, 1, 3)+t(1, 1, -1)\}$$

con cada uno de los planos de coordenadas.

5. Determínese el punto donde la recta $\mathscr{L}$ que pasa por el punto $(1, 3, 1)$ y es ortogonal al plano $\mathscr{P} : 3x-2y+5z = 15$ intersecta a $\mathscr{P}$.

6. Una partícula comienza a moverse en el punto $(15, -22, 10)$ y se mueve con una velocidad constante $(1, 1, 1)$. ¿Cuánto tarda la partícula en alcanzar al plano $x+10y+4z = -15$?

7. ¿En qué dirección debería moverse la partícula del problema 6 para alcanzar el plano en tiempo mínimo?, si el valor absoluto de la velocidad es el mismo que en el problema 6, ¿cuál es el tiempo mínimo?

8. Demuéstrese que los planos

$$\mathscr{P}_1 = \{(2, 0, 4)+u(1, 7, 3)+v(-3, 8, 0)\}$$

y

$$\mathscr{P}_2 = \{(3, 2, 3)+s(4, -1, 3)+t(9, 5, 9)\}$$

son paralelos. Encuéntrese la distancia entre $\mathscr{P}_1$ y $\mathscr{P}_2$ si definimos la distancia entre planos paralelos se define como la distancia de un punto cualquiera de un plano al otro plano.

9. Sea $\mathscr{L}$ la intersección de los planos con ecuaciones $3x+y-4z = 5$ y $2x+3y-z = 4$. Si $\mathscr{P}$ es el plano con ecuación $x-2y+3z = 1$, encuéntrese $\mathscr{L} \cap \mathscr{P}$.

10. Encuéntrese una ecuación del plano que contiene el punto $(1, 2, -3)$ y la recta $\mathscr{L} = \{(1, 1, 1)+t(5, -2, 3)\}$.

10. BASES

Hemos demostrado que cualquier vector $\mathbf{a} \in V_3$ puede expresarse en una forma única como una combinación lineal de los vectores unitarios

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1).$$

En realidad,

$$\mathbf{a} = (a_1, a_2, a_3) = a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}.$$

Así pues, los vectores **i**, **j**, **k** tienen la propiedad de que todo vector de V_3 puede expresarse como una combinación lineal de estos vectores. El conjunto de vectores **i**, **j**, **k** no es el único conjunto de vectores que tiene esta propiedad. Demostraremos que cualesquier tres vectores linealmente independientes tienen esta propiedad.

10.1 Teorema. *Si* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V_3$ *son linealmente independientes, entonces para cada punto* $\mathbf{P} \in \mathrm{R}^3$ *existen números reales únicos* u, v, t *tales que*

$$\mathbf{P} = u\mathbf{a} + v\mathbf{b} + t\mathbf{c}.$$

PRUEBA. De acuerdo con el problema 5, pág. 66, como **a**, **b**, **c** son linealmente independientes, si existen números u, v, t, que satisfacen

$$\mathbf{P} = u\mathbf{a} + v\mathbf{b} + t\mathbf{c},$$

tales números son únicos. Que tales números existan es una consecuencia del teorema 9.2. Sea $\mathscr{P} = \{u\mathbf{a} + v\mathbf{b}\}$ y $\mathscr{L} = \{\mathbf{P} + t\mathbf{c}\}$. Como **a**, **b** y **c** son linealmente independientes, $\mathscr{L}$ no es paralela a $\mathscr{P}$. Luego si $\mathbf{P}_1$ es el punto de intersección de $\mathscr{P}$ y $\mathscr{L}$ hay números $u, v, t_1 \in \mathrm{R}$ tales que

$$\mathbf{P}_1 = u\mathbf{a} + v\mathbf{b} = \mathbf{P} + t_1\mathbf{c}.$$

Por tanto

$$\mathbf{P} = u\mathbf{a} + v\mathbf{b} + t\mathbf{c}$$

donde $t = -t_1$.

Nota. Como una consecuencia del teorema 10.1, vemos que cualquier conjunto de cuatro o más vectores en V_3 es linealmente dependiente.

10.2 Definición. *Un conjunto* $\{a_1, \ldots, a_k\}$ *de vectores en* V_n *se dice que es una base de* V_n *si*

i) $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ *es linealmente independiente*

y

ii) *todo vector de* V_n *puede expresarse como una combinación lineal de* $\mathbf{a}_1, \ldots, \mathbf{a}_k$.

Si todo vector de V_n puede expresarse como una combinación lineal de $\mathbf{a}_1, \ldots, \mathbf{a}_k$, entonces el conjunto $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ se dice que *genera* V_n.

10.3 Corolario. *Todo conjunto de tres vectores linealmente independientes de* V_3 *es una base de* V_3.

PRUEBA. Sean **a**, **b**, **c** tres vectores linealmente independientes de V_3. Solo necesitamos demostrar que **a**, **b**, **c** generan V_3. Sea **d** un vector arbitrario de V_3. Sea $\mathbf{P} \in \mathrm{R}^3$ tal que $\mathbf{P} = \mathbf{P} - \mathbf{O} = \mathbf{d}$. Entonces, de acuerdo con el teorema 10.1, existen números reales únicos u, v, t tales que

$$\mathbf{d} = \mathbf{P} = u\mathbf{a} + v\mathbf{b} + t\mathbf{c}.$$

10.4 Corolario. *Si*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0$$

entonces el sistema de ecuaciones linales

$$a_1 x + b_1 y + c_1 z = d_1$$
$$a_2 x + b_2 y + c_2 z = d_2$$
$$a_3 x + b_3 y + c_3 z = d_3$$

tiene una solución única.

PRUEBA. El sistema de tres ecuaciones lineales es equivalente a la ecuación vectorial

10.5 $$\mathbf{a}x + \mathbf{b}y + \mathbf{c}z = \mathbf{d}$$

donde $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$, $\mathbf{c} = (c_1, c_2, c_3)$ y $\mathbf{d} = (d_1, d_2, d_3)$. Como

$$[\mathbf{abc}] = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0,$$

a, **b**, **c** son linealmente independientes (teorema 6.4, pág. 64). De donde se sigue, de acuerdo con el corolario 10.3, que la ecuación 10.5 tiene una solución. La independencia lineal de **a**, **b**, **c** implica la unicidad de la solución (problema 5, pág. 66).

Hallando los productos escalares de la ecuación 10.5 por $\mathbf{b}\times\mathbf{c}$, $\mathbf{c}\times\mathbf{a}$, y $\mathbf{a}\times\mathbf{b}$, sucesivamente, obtenemos

$$[\mathbf{abc}]x = [\mathbf{dbc}], \quad [\mathbf{abc}]y = [\mathbf{adc}], \quad [\mathbf{abc}]z = [\mathbf{abd}]$$

o bien

$$x = \frac{[\mathbf{dbc}]}{[\mathbf{abc}]} = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}, \qquad y = \frac{[\mathbf{adc}]}{[\mathbf{abc}]} = \frac{\begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}},$$

$$z = \frac{[\mathbf{abd}]}{[\mathbf{abc}]} = \frac{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}$$

Nota. La fórmula para la solución de corolario 10.4 es el caso especial para tres ecuaciones lineales con tres incógnitas de la *regla de Cramer:* si el determinante de coeficientes es distinto de cero en un sistema de n ecuaciones lineales con n incógnitas, entonces cada una de las incógnitas puede expresarse como el cociente de dos determinantes —el denominador es el determinante de los coeficientes y el numerador para la incógnita j-ésima es el determinante obtenido cuando en el determinante de los coeficientes se reemplaza la j-ésima columna por la columna de las constantes. Aunque la regla de Cramer nos da una solución formal a un sistema de n ecuaciones lineales con n incógnitas cuando el determinante de los coeficientes es distinto de cero, no nos proporciona sin embargo un método práctico de solución, excepto para los casos en que n es pequeña. La evaluación de un determinante de orden n requiere $(n-1)\times n!$ multiplicaciones, luego resolver un sistema de n ecuaciones por la regla de Cramer requeriría $(n-1)\times(n+1)!$ multiplicaciones. La reducción de Gauss-Jordan, de la que damos una muestra en la solución del ejemplo 10.7 que después del próximo corolario exponemos, requiere solamente $\frac{1}{3}(n^3+3n^2-n)$ multiplicaciones. Por ejemplo, para $n = 10$, la regla de Cramer requiere casi 359 millones de multiplicaciones, mientras que la reducción de Gauss-Jordan requiere solamente 430. La mayor parte de los métodos prácticos de resolución de tales sistemas son variaciones de la reducción de Gauss-Jordan.

10.6 Corolario. *Tres planos cuyas normales son linealmente independientes se intersectan en un y sólo en un punto.*

PRUEBA. Esta es la interpretación geométrica del anterior corolario cuando las ecuaciones que se consideran son las ecuaciones de tres planos. Un punto de intersección de los planos corresponde a una solución del sistema de ecuaciones.

10.7 Ejemplo. Encuéntrense todos los puntos de intersección de los planos

$$\begin{aligned} 2x+y-3z &= 4 \\ 5x+4y+7z &= 2 \\ x+y+2z &= -5. \end{aligned}$$

SOLUCIÓN. Escribiendo primero la última ecuación, resolviéndola para x, y sustituyendo en las otras dos ecuaciones, tenemos:

$$\begin{cases} x = y-2z-5 \\ 3y-7z = 14 \\ 9y-3z = 27 \end{cases}$$

$$\Leftrightarrow \begin{cases} x = y - 2z - 5 \\ y = \frac{7}{3}z + \frac{14}{3} \\ z = -\frac{5}{6} \end{cases}$$

$$\Leftrightarrow \begin{cases} x = -\frac{11}{18} \\ y = \frac{49}{18} \\ z = -\frac{5}{6}. \end{cases}$$

El punto de intersección es $(-\frac{11}{18}, \frac{49}{18}, -\frac{5}{6})$. De nuevo señalamos que aunque no es lógicamente necesario comprobar la solución, sin embargo, es prudente hacerlo.

10.8 Ejemplo. Si tres vectores no nulos **a**, **b**, **c** en V_3 son mutuamente ortogonales, pruébese que forman una base de V_3. Exprésese un vector $\mathbf{d} \in V_3$ como una combinación lineal de **a**, **b** y **c**.

SOLUCIÓN. Para que **a**, **b**, **c** constituyan una base de V_3 han de ser linealmente independientes. Es decir, debemos demostrar que la única solución de la ecuación

10.9 $$\mathbf{0} = u\mathbf{a} + v\mathbf{b} + t\mathbf{c}$$

es $u = v = t = 0$. Tomando el producto escalar de ambos miembros de la ecuación 10.9 por **a**, tenemos

$$\mathbf{a} \cdot \mathbf{0} = u\mathbf{a} \cdot \mathbf{a} + v\mathbf{a} \cdot \mathbf{b} + t\mathbf{a} \cdot \mathbf{c}.$$

Como **a** es ortogonal tanto a **b** como a **c**, obtenemos $u = 0$. Análogamente, tomando el producto escalar con **b** y con **c**, obtenemos $v = 0$ y $t = 0$. De donde **a**, **b**, **c** son linealmente independientes, luego, según corolario 10.3, forman una base de V_3.

Como **a**, **b**, **c** forman una base de V_3, todo vector $\mathbf{d} \in V_3$ puede expresarse como una combinación lineal

$$\mathbf{d} = u\mathbf{a} + v\mathbf{b} + t\mathbf{c}.$$

Tomando el producto escalar con **a**, **b**, **c** sucesivamente, como **a**, **b**, **c** son mutuamente ortogonales, obtenemos

$$\mathbf{a} \cdot \mathbf{d} = u\mathbf{a} \cdot \mathbf{a}, \quad \mathbf{b} \cdot \mathbf{d} = v\mathbf{b} \cdot \mathbf{b}, \quad \mathbf{c} \cdot \mathbf{d} = t\mathbf{c} \cdot \mathbf{c}$$

o bien

$$u = \frac{\mathbf{a} \cdot \mathbf{d}}{|\mathbf{a}|^2}, \quad v = \frac{\mathbf{b} \cdot \mathbf{d}}{|\mathbf{b}|^2}, \quad t = \frac{\mathbf{c} \cdot \mathbf{d}}{|\mathbf{c}|^2}.$$

Problemas

1. Demuéstrese que los vectores **a**, **b**, **c** son mutuamente ortogonales y exprésese **d** como una combinación lineal de **a**, **b**, **c**.

a) $\mathbf{a} = (1, 0, 0)$, $\mathbf{b} = (0, 1, 1)$, $\mathbf{c} = (0, -1, 1)$, $\mathbf{d} = (3, 4, -2)$
b) $\mathbf{a} = (1, 1, 0)$, $\mathbf{b} = (0, 0, 3)$, $\mathbf{c} = (1, -1, 0)$, $\mathbf{d} = (-2, 5, 1)$
c) $\mathbf{a} = (1, 2, 1)$, $\mathbf{b} = (-1, 2, -3)$, $\mathbf{c} = (-4, 1, 2)$, $\mathbf{d} = (1, 3, 5)$
d) $\mathbf{a} = (1, 1, 1)$, $\mathbf{b} = (2, -3, 1)$, $\mathbf{c} = (4, 1, -5)$, $\mathbf{d} = (2, 6, -7)$.

2. Sean **a**, **b**, $\mathbf{c} \in V_3$ linealmente independientes. Demuéstrese que **a**, $\mathbf{b}_1$, $\mathbf{c}_1$, donde

$$\mathbf{b}_1 = \mathbf{b} - \mathbf{Proy_a b} \quad \text{y} \quad \mathbf{c}_1 = \mathbf{c} - \mathbf{Proy_a c} - \mathbf{Proy_{b_1} c}$$

forman una base de vectores mutuamente ortogonales de V_3. La interpretación geométrica de $\mathbf{b}_1$ y $\mathbf{c}_1$ se muestra en la figura 11. Este proceso se conoce como el proceso de ortogonalización de Gram-Schmidt.

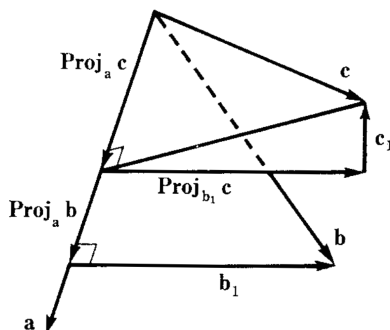


FIGURA 11

3. Usese el método del problema 2 para obtener una base ortogonal de V_3, partiendo de los vectores **a**, **b**, **c** en cada uno de los siguientes casos:

a) $\mathbf{a} = (1, 2, 1)$, $\mathbf{b} = (2, -3, 2)$, $\mathbf{c} = (2, -1, 1)$
b) $\mathbf{a} = (1, 1, 1)$, $\mathbf{b} = (-2, 3, 1)$, $\mathbf{c} = (1, 2, -1)$
c) $\mathbf{a} = (1, 0, 0)$, $\mathbf{b} = (0, 1, 0)$, $\mathbf{c} = (0, 0, 1)$.

4. Exprésese el vector **d** como una combinación lineal de **a**, **b** y **c**.

a) $\mathbf{a} = (1, 2, 1)$, $\mathbf{b} = (2, -3, 2)$, $\mathbf{c} = (2, -1, 1)$, $\mathbf{d} = (3, 4, -2)$
b) $\mathbf{a} = (1, 1, 1)$, $\mathbf{b} = (-2, 3, 1)$, $\mathbf{c} = (1, 2, -1)$, $\mathbf{d} = (5, -7, 2)$.

5. Encuéntrense todos los puntos de intersección de los planos

a)
$$\begin{aligned} 3x+y+z &= 5 \\ 3x+y+5z &= 7 \\ x-y+3z &= 3 \end{aligned}$$

b)
$$\begin{aligned} 5x+y-z &= 6 \\ -2x+y-4z &= 10 \\ x-3y+z &= 8. \end{aligned}$$

6. Demuéstrese que el conjunto de vectores $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, donde $\mathbf{e}_1 = (1, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \mathbf{e}_n = (0, \ldots, 0, 1)$, es una base de V_n.

11. COORDENADAS CILÍNDRICAS Y ESFÉRICAS

En el espacio bidimensional ocurre a menudo que es conveniente expresar la ecuación de una curva en coordenadas polares o en algún otro sistema de coordenadas distinto del de coordenadas rectangulares. Análogamente, en el espacio tridimensional a menudo son útiles otros sistemas de coordenadas distintos del de coordenadas rectangulares. Las coordenadas de uso más común en el espacio tridimensional, aparte de las rectangulares, son las coordenadas cilíndricas y las coordenadas esféricas.

Los significados geométricos de las coordenadas cilíndricas y esféricas se muestran en las figuras 12 y 13. Denotamos a las coordenadas cilíndricas por (r, θ, z) y a las coordenadas esféricas por (ρ, θ, φ). Si (r, θ, z) son las coordenadas cilíndricas de un punto $\mathbf{P}$ en R^3, entonces

11.1 $$\mathbf{P} = (r \cos \theta, r \operatorname{sen} \theta, z).$$

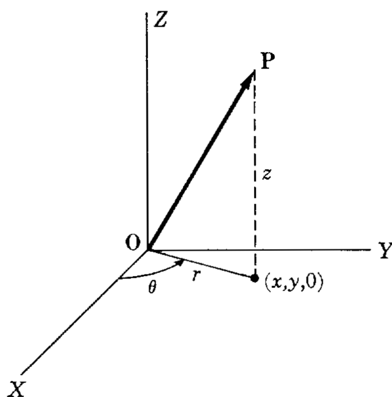


FIGURA 12 $\mathbf{P} = (x,y,z)$

Así pues, si $\mathbf{P} = (x, y, z)$, la relación entre las coordenadas cartesianas x, y, z de $\mathbf{P}$ y las coordenadas cilíndricas de $\mathbf{P}$ es

11.2 $$x = r \cos \theta$$
$$y = r \operatorname{sen} \theta$$
$$z = z.$$

Las coordenadas cilíndricas r y θ son las coordenadas polares del punto $(x, y, 0)$ en el plano XY: $(x, y, 0)$ es la proyección ortogonal de $\mathbf{P}$ sobre

el plano XY. La coordenada cilíndrica z de **P** es la altura de **P** sobre el plano XY. Así pues, según 11.1 o su equivalente 11.2, cada conjunto (r, θ, z) de coordenadas cilíndricas determina un punto único de R^3. Recíprocamente, a cada punto **P** de R^3 se le pueden asignar coordenadas cilíndricas, es decir, dado **P** pueden encontrarse números (r, θ, z) que satisfagan 11.1. Esta asignación de coordenadas cilíndricas no es única. Excluido el eje Z, las restricciones $r > 0$ y $0 \leqslant \theta < 2\pi$ hacen única la asignación.

Si (ρ, θ, φ) son coordenadas esféricas (figura 13) de un punto **P**, entonces

11.3 $$\mathbf{P} = (\rho \operatorname{sen} \varphi \cos \theta, \rho \operatorname{sen} \varphi \operatorname{sen} \theta, \rho \cos \varphi).$$

Cada conjunto (ρ, θ, φ) de coordenadas esféricas determina un punto único **P** de R^3. La coordenada esférica θ es la misma que la coordenada cilíndrica θ y suele llamarse *longitud* o *acimut* de **P**. La coordenada esférica φ es el ángulo entre la dirección positiva del eje Z y el radio vector **P**. A este ángulo φ se le llama *colatitud* de **P** $\left(\frac{\pi}{2} - \varphi \text{ se llama } latitud \text{ de } \mathbf{P}\right)$. La coordenada esférica ρ es la distancia de **P** al origen (si $\rho > 0$). La relación entre coordenadas esféricas y coordenadas cartesianas es

11.4 $$\begin{aligned} x &= \rho \operatorname{sen} \varphi \cos \theta \\ y &= \rho \operatorname{sen} \varphi \operatorname{sen} \theta \\ z &= \rho \cos \varphi . \end{aligned}$$

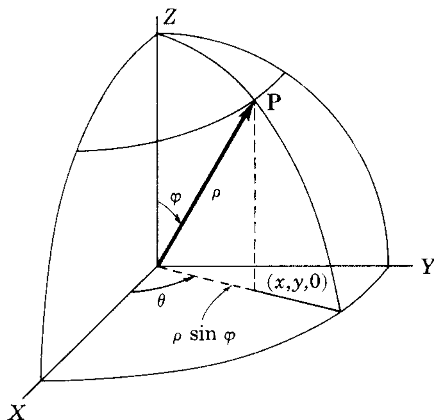


FIGURA 13 $\mathbf{P} = (x,y,z)$

Es claro que todo punto de R^3 tiene conjuntos de coordenadas esféricas. De nuevo aquí, si excluimos los puntos del eje Z, las restricciones $\rho > 0$, $0 \leqslant \theta < 2\pi$, y $0 < \varphi < \pi$ hacen única la asignación de coordenadas esféricas.

Nota. Las notaciones para las coordenadas esféricas no son universales. Muchos autores usan θ para la colatitud, a la que nosotros hemos denotado por φ, y φ para la longitud a la que nosotros denotamos por θ. Cuando se hace esto, θ no tiene la misma significación en coordenadas cilíndricas que en coordenadas esféricas, mientras que con la elección de notaciones que nosotros hemos hecho, θ significa lo mismo en las coordenadas cilíndricas que en las esféricas. Otra dificultad que aparece en la notación más común es que la misma letra, generalmente r, se usa para denotar tanto la distancia al eje en las coordenadas cilíndricas (r en nuestra notación) como la distancia al origen en las coordenadas esféricas (ρ en nuestra notación). Algunos autores, particularmente en el campo de la mecánica de fluidos, usan ω tanto en las coordenadas cilíndricas como en las esféricas para representar la longitud que aquí hemos denotado por θ. En este caso, θ se usa para denotar la colatitud, que aquí aparece representada por φ.

11.5 Definición. *Un* ***cilindro*** *(circular recto) es un conjunto de puntos equidistantes de una recta fija a la que se llama eje del cilindro.*

La distancia de un punto $\mathbf{P} = (x, y, z)$ al eje Z es $\sqrt{x^2+y^2}$. Vemos, pues, que el conjunto

11.6 $$\mathscr{C} = \{(x, y, z) \mid x^2+y^2 = a^2\}$$

es un cilindro circular de radio a cuyo eje es el eje Z. La ecuación $x^2+y^2 = a$ se dice que es una ecuación (en coordenadas cartesianas) del cilindro $\mathscr{C}$. En términos de coordenadas cilíndricas, $\mathscr{C} = \{(r \cos \theta, r \operatorname{sen} \theta, z) \mid r = a\}$ es el mismo cilindro, y $r = a$ se llama ecuación del cilindro $\mathscr{C}$ en coordenadas cilíndricas. Nótese también que podemos escribir[1]

$$\mathscr{C} = \{(a \cos u, a \operatorname{sen} u, v) \mid u \in [0, 2\pi], v \in \langle -\infty, \infty\rangle\}.$$

Las ecuaciones

11.7 $$\begin{aligned} x &= a \cos u \\ y &= a \operatorname{sen} u \\ z &= v, \quad u \in [0, 2\pi], \quad v \in \langle -\infty, \infty\rangle, \end{aligned}$$

se llaman ecuaciones paramétricas del cilindro $\mathscr{C}$.

11.8 Definición. *La* ***esfera*** $\mathscr{S}(\mathbf{C}; a)$ *de radio* a *y centro en el punto* $\mathbf{C} = (c_1, c_2, c_3)$ *es el conjunto de todos los puntos cuya distancia de* $\mathbf{C}$ *es* a*; es decir,*

$$\begin{aligned} \mathscr{S}(\mathbf{C}; a) &= \{\mathbf{P} \mid |\mathbf{P}-\mathbf{C}| = a\} \\ &= \{(x, y, z) \mid (x-c_1)^2+(y-c_2)^2+(z-c_3)^2 = a^2\}. \end{aligned}$$

[1] Al intervalo cerrado determinado por los números reales a y b ($a \leqslant b$) le representamos por $[a, b]$ y al intervalo abierto correspondiente por $\langle a, b\rangle$; es decir, $[a, b] = \{x \mid a \leqslant x \leqslant b\}$ y $\langle a, b\rangle = \{x \mid a < x < b\}$.

La ecuación

11.9 $$(x-c_1)^2+(y-c_2)^2+(z-c_3)^2 = a^2$$

se llama ecuación en coordenadas cartesianas de la esfera $\mathscr{S}(\mathbf{C}; a)$.[1] La esfera $\mathscr{S}(\mathbf{O}; a)$ de radio a y centro en el origen, se puede expresar en coordenadas esféricas en la forma

11.10 $$\mathscr{S}(\mathbf{O}; a) = \{(\rho \operatorname{sen} \varphi \cos \theta, \rho \operatorname{sen} \varphi \operatorname{sen} \theta, \rho \cos \varphi) \mid \rho = a\}.$$

La ecuación $\rho = a$ se llama ecuación en coordenadas esféricas de la esfera $\mathscr{S}(\mathbf{O}; a)$. Podemos también escribir

$$\mathscr{S}(\mathbf{O}; a) = \{(a \operatorname{sen} u \cos v, a \operatorname{sen} u \operatorname{sen} v, a \cos u \mid u \in [0, \pi], v \in [0, 2\pi]\}.$$

Las ecuaciones

11.11 $$\begin{aligned} x &= a \operatorname{sen} u \cos v \\ y &= a \operatorname{sen} u \operatorname{sen} v \\ z &= a \cos u, \quad u \in [0, \pi], \quad v \in [0, 2\pi] \end{aligned}$$

se llaman ecuaciones paramétricas de la esfera $\mathscr{S}(\mathbf{O}; a)$.

Problemas

1. Determínense las coordenadas cartesianas rectangulares de los puntos cuyas coordenadas cilíndricas (r, θ, z) son:

a) $\left(1, \frac{\pi}{2}, 0\right)$ *b*) $\left(4, -\frac{\pi}{3}, 0\right)$ *c*) $\left(3, \frac{3\pi}{2}, 10\right)$

d) $(1, 1, 1)$ *e*) $\left(-1, \frac{\pi}{2}, -3\right)$ *f*) $(6, 28^\circ, 7)$.

2. Determínense las coordenadas cartesianas de los puntos cuyas coordenadas esféricas (ρ, θ, φ) son

a) $\left(1, 0, \frac{\pi}{2}\right)$ *b*) $(1, \theta, 0)$ *c*) $\left(1, \frac{\pi}{2}, \frac{\pi}{4}\right)$

d) $(1, 1, 1)$ *e*) $(13, 120^\circ, 30^\circ)$ *f*) $(-1, -1, -1)$.

3. Asígnense coordenadas cilíndricas y esféricas a los puntos cuyas coordenadas cartesianas son:

a) $(1, 0, 0)$ *b*) $(1, 1, 0)$ *c*) $(0, 0, 1)$ *d*) $(0, 3, -1)$
e) $(0, 0, 0)$ *f*) $(1, 1, 1)$ *g*) $(32, -25, 18)$

[1] Nótese que estamos llamando esfera a lo que en la literatura matemática castellana es habitual llamar superficie esférica. [N. del T.]

4. Pruébese que la hélice cilíndrica

$$\mathscr{C} : \mathbf{f}(t) = (a \cos \omega t, a \operatorname{sen} \omega t, bt), \quad t \in \langle -\infty, \infty \rangle,$$

se encuentra sobre un cilindro.

5. Descríbase la superficie cuya ecuación en coordenadas cilíndricas es:

a) $r = \text{const}$ *b*) $\theta = \text{const}$ *c*) $z = \text{const}$.

6. Descríbase la superficie cuya ecuación en coordenadas esféricas es:

a) $\rho = \text{const}$ *b*) $\theta = \text{const}$ *c*) $\varphi = \text{const}$.

12. ESPACIOS EUCLIDIANOS *n*-DIMENSIONALES

En esta sección definiremos el espacio euclidiano n-dimensional. Tomaremos como modelo el espacio euclidiano tridimensional. Hemos visto que en el espacio tridimensional tres vectores linealmente independientes generan el espacio total (corolario 10.3, pág. 80). Dos vectores linealmente independientes $\mathbf{a}, \mathbf{b} \in V_3$ determinan un plano $\{u\mathbf{a}+v\mathbf{b} \mid u, v \in \mathrm{R}\}$ que pasa por el origen. Un plano así se llama subespacio bidimensional de R^3 y cada punto del subespacio está únicamente determinado por los dos parámetros u, v. Todo plano en R^3 resulta de aplicar una traslación a tal subespacio, $\{\mathbf{P}_0+u\mathbf{a}+v\mathbf{b} \mid u, v \in \mathrm{R}\}$, de tal subespacio. Un vector linealmente independiente, es decir, un vector distinto de cero, determina una recta $\{t\mathbf{a} \mid t \in \mathrm{R}\}$ que pasa por el origen. Tal recta se llama subespacio unidimensional de R^3 y cada punto de este subespacio está determinado inequívocamente por el parámetro único t. Toda recta en R^3 es una traslación, $\{\mathbf{P}_0+t\mathbf{a} \mid t \in \mathrm{R}\}$, de tal subespacio. En un espacio n-dimensional, podemos tener k vectores linealmente independientes $\mathbf{a}_1, \ldots, \mathbf{a}_k$ donde $k = 1, \ldots, n$. También en este caso cuando $k = n$, los vectores $\mathbf{a}_1, \ldots, \mathbf{a}_n$ generan el espacio total y cuando $k < n$ generan un subespacio k-dimensional. Tales subespacios $(t_1\mathbf{a}_1 + \ldots + t_k\mathbf{a}_k \mid t_1, \ldots, t_k \in \mathrm{R}\}$ y sus traslaciones, conjuntos de la forma $\{\mathbf{P}_0 + t_1\mathbf{a}_1 + \ldots + t_k\mathbf{a}_k \mid t_1, \ldots, t_k \in \mathrm{R}\}$, se llaman planos k-dimensionales en el espacio n-dimensional. El plano unidimensional se llama usualmente recta.

12.1 Definición. *El **espacio** "analítico" n-**dimensional euclidiano**, denotado por* R^n *es el espacio vectorial n-dimensional* V_n *donde*:

i) *los elementos* $\mathbf{x} = (x_1, \ldots, x_n)$ *de* V_n *son los **puntos** de* R^n,

ii) *un conjunto* $\mathscr{P}$ *de puntos de* R^n *se llama **plano** k-**dimensional** o hiperplano en* R^n $(k = 1, \ldots, n-1)$ *si hay un punto* $\mathbf{P}_0 \in \mathrm{R}^n$ *y k vectores linealmente independientes* $\mathbf{a}_1, \ldots, \mathbf{a}_k \in V_n$ *tales que*

12.2 $$\mathscr{P} = \{\mathbf{P}_0 + t_1\mathbf{a}_1 + \ldots + t_k\mathbf{a}_k \mid t_1, \ldots, t_k \in \mathrm{R}\};$$

iii) *la **distancia**, escrita* $d(\mathbf{P}, \mathbf{Q})$, *desde el punto* $\mathbf{P} = (x_1, \ldots, x_n)$ *al punto* $\mathbf{Q} = (y_1, \ldots, y_n)$ *en* R^n *es la longitud del vector* $\mathbf{Q}-\mathbf{P}$, *es decir*

$$d(\mathbf{P}, \mathbf{Q}) = |\mathbf{Q}-\mathbf{P}| = [(y_1-x_1)^2 + \ldots + (y_n-x_n)^2]^{1/2}.$$

Un plano unidimensional

$$\mathscr{L} = \{\mathbf{P}_0 + t\mathbf{a} \mid t \in \mathrm{R}\} \quad (\mathbf{a} \neq \mathbf{0})$$

se llama *recta*.

La ecuación

$$\mathbf{P} = \mathbf{P}_0 + t_1 \mathbf{a}_1 + \ldots + t_k \mathbf{a}_k$$

se llama *ecuación vectorial* de un plano y las ecuaciones de las componentes se llaman *ecuaciones paramétricas* del plano.

Decimos que un vector $\mathbf{b}$ es paralelo al plano $\mathscr{P} = \{\mathbf{P}_0 + t_1 \mathbf{a}_1 + \ldots t_k \mathbf{a}_k \mid t_1, \ldots, t_k \in \mathrm{R}\}$ si $\mathbf{b} = r_1 \mathbf{a}_1 + \ldots + r_k \mathbf{a}_k$ para algunos $r_1, \ldots, r_k \in \mathrm{R}$.

12.3 Ejemplo. Determínese una recta (plano unidimensional) que pase por los puntos $\mathbf{P}_0 = (1, 4, -2, 3)$ y $\mathbf{P}_1 = (2, 0, 1, 0)$ de R^4.

SOLUCIÓN. Sea $\mathbf{a} = \mathbf{P}_1 - \mathbf{P}_0$. Entonces

$$\mathscr{L} = \{\mathbf{P}_0 + t(\mathbf{P}_1 - \mathbf{P}_0) \mid t \in \mathrm{R}\}$$

es una recta que contiene a $\mathbf{P}_0$ y a $\mathbf{P}_1$: $\mathbf{P}_0$ corresponde a $t = 0$ y $\mathbf{P}_1$ a $t = 1$. Luego

$$\mathscr{L} = \{(1, 4, -2, 3) + t(1, -4, 3, -3)\}$$

es una recta (1, 4, −2, 3) y (2, 0, 1, 0).

12.4 Ejemplo. Determínese un plano que pase por los puntos $\mathbf{P}_0 = (2, 3, -2, 3)$, $\mathbf{P}_1 = (3, 2, 1, 0)$, $\mathbf{P}_2 = (2, 1, 0, 0)$ y $\mathbf{P}_3 = (2, 0, 2, 0)$.

SOLUCIÓN. Sean

$$\begin{aligned} \mathbf{a}_1 &= \mathbf{P}_1 - \mathbf{P}_0 = (1, -1, 3, -3), \\ \mathbf{a}_2 &= \mathbf{P}_2 - \mathbf{P}_0 = (0, -2, 2, -3), \\ \mathbf{a}_3 &= \mathbf{P}_3 - \mathbf{P}_0 = (0, -3, 4, -3). \end{aligned}$$

Con el fin de determinar la dimensión del plano necesitamos conocer cuántos de los anteriores tres vectores son linealmente independientes. Los tres vectores son linealmente dependientes si existen números r_1, r_2, r_3, no todos ceros, tales que

$$r_1 \mathbf{a}_1 + r_2 \mathbf{a}_2 + r_3 \mathbf{a}_3 = \mathbf{0}.$$

Esta ecuación vectorial es equivalente a las cuatro ecuaciones componentes:

$$\begin{aligned} r_1 \qquad\qquad &= 0 \\ -r_1-2r_2-3r_3 &= 0 \\ 3r_1+2r_2+4r_3 &= 0 \\ -3r_1-3r_2-3r_3 &= 0. \end{aligned}$$

Se encuentra que la única solución a estas ecuaciones es $r_1=r_2=r_3=0$. Luego los tres vectores a_1, a_2, a_3 son linealmente independientes. Por tanto

$$\begin{aligned} \mathscr{P} &= \{\mathbf{P}_0+t_1\mathbf{a}_1+t_2\mathbf{a}_2+t_3\mathbf{a}_3\} \\ &= \{(2, 3, -2, 3)+t_1(1, -1, 3, -3)+t_2(0, -2, 2, -3)+t_3(0, -3, 4, -3)\} \end{aligned}$$

es un plano tridimensional en R^4 que pasa por los puntos $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{P}_2$ y $\mathbf{P}_3$: $\mathbf{P}_0$ corresponde a $t_1=t_2=t_3=0$; $\mathbf{P}_1$ corresponde a $t_1=1$, $t_2=t_3=0$; $\mathbf{P}_2$ corresponde a $t_2=1$, $t_1=t_3=0$; y $\mathbf{P}_3$ corresponde a $t_3=1$, $t_1=t_2=0$.

12.5 Ejemplo. Determínese la intersección de la recta $\mathscr{L}$ del ejemplo 12.3 y el plano $\mathscr{P}$ del ejemplo 12.4.

SOLUCIÓN. Supongamos $\mathbf{P}\in\mathscr{L}\cap\mathscr{P}$. Entonces $\mathbf{P}\in\mathscr{L}$ implica

$$\mathbf{P} = (1, 4, -2, 3)+t(1, -4, 3, -3) \text{ para algún } t\in\mathrm{R}$$

y $\mathbf{P}\in\mathscr{P}$ implica

$$\mathbf{P} = (2, 3, -2, 3)+t_1(1, -1, 3, -3)+t_2(0, -2, 2, -3)+t_3(0, -3, 4, -3) \quad \text{para algunos } t_1, t_2, t_3\in\mathrm{R}.$$

Por tanto $\mathbf{P}$ pertenece a la intersección si y sólo si estas dos expresiones son iguales; es decir, si y sólo si

$$\begin{aligned} 1+t &= 2+t_1 \\ 4-4t &= 3-t_1-2t_2-3t_3 \\ -2+3t &= -2+3t_1+2t_2+4t_3 \\ 3-3t &= 3-3t_1-3t_2-3t_3. \end{aligned}$$

Resolviendo estas ecuaciones, encontramos $t=\frac{5}{6}$, $t_1=-\frac{1}{6}$, $t_2=\frac{1}{2}$ y $t_3=\frac{1}{2}$. Por tanto

$$\mathbf{P} = (1, 4, -2, 3)+\tfrac{5}{6}(1, -4, 3, -3) = \tfrac{1}{6}(11, 4, 3, 3)$$

es *el* punto de intersección de $\mathscr{L}$ y $\mathscr{P}$.

COMPROBACIÓN.

$$(2, 3, -2, 3)-\tfrac{1}{6}(1, -1, 3, -3)+\tfrac{1}{2}(0, -2, 2, -3)+\tfrac{1}{2}(0, -3, 4, -3) = \tfrac{1}{6}(11, 4, 3, 3).$$

12.6 Definición. *Dos planos en* R^n,

$$\mathscr{P}_1 = \{\mathbf{P}_1+s_1\mathbf{a}_1+\ldots+s_j\mathbf{a}_j \mid s_1, \ldots, s_j\in\mathrm{R}\}$$

y

$$\mathscr{P}_2 = \{\mathbf{P}_2 + t_1 \mathbf{b}_1 + \ldots + t_k \mathbf{b}_k \mid t_1, \ldots, t_k \in \mathrm{R}\}$$

donde $j \leqslant k < n$, son ***paralelos*** *si cada uno de los conjuntos de $k+1$ vectores $\{\mathbf{a}_i, \mathbf{b}_1, \ldots, \mathbf{b}_k\}$ $(i = 1, \ldots, j)$ es linealmente dependiente; es decir, si cada uno de los vectores $\mathbf{a}_i$ $(i = 1, \ldots, j)$ es paralelo al plano $\mathscr{P}_2$.*

Mostramos ahora que si dos planos son paralelos, entonces o un plano es un subconjunto del otro o su intersección es vacía. Este teorema contiene el corolario 3.4 y las partes de los teoremas 8.2 y 9.2 concernientes a las rectas y planos paralelos como casos particulares.

12.7 Teorema. *Si dos planos en* R^n

$$\mathscr{P}_1 = \{\mathbf{P}_1 + s_1 \mathbf{a}_1 + \ldots + s_j \mathbf{a}_j \mid s_1, \ldots, s_j \in \mathrm{R}\}$$

y

$$\mathscr{P}_2 = \{\mathbf{P}_2 + t_1 \mathbf{b}_1 + \ldots + t_k \mathbf{b}_k \mid t_1, \ldots, t_k \in \mathrm{R}\}$$

donde $j \leqslant k < n$, son paralelos, entonces $\mathscr{P}_1 \subset \mathscr{P}_2$ o $\mathscr{P}_1 \cap \mathscr{P}_2 = \varnothing$.

PRUEBA. Como $\mathbf{b}_1, \ldots, \mathbf{b}_k$ son linealmente independientes, $\mathbf{a}_i, \mathbf{b}_1, \ldots, \mathbf{b}_k$ linealmente dependientes implica que existen números λ_{im} tales que

12.8 $$\mathbf{a}_i = \sum_{m=1}^{k} \lambda_{im} \mathbf{b}_m \quad (i = 1, \ldots, j).$$

Supongamos $\mathbf{P}_0 \in \mathscr{P}_1 \cap \mathscr{P}_2$. Entonces hay números $s_1', \ldots, s_j', t_1', \ldots, t_k'$ tales que

$$\mathbf{P}_0 = \mathbf{P}_1 + s_1' \mathbf{a}_1 + \ldots + s_j' \mathbf{a}_j = \mathbf{P}_2 + t_1' \mathbf{b}_1 + \ldots + t_k' \mathbf{b}_k.$$

Por tanto

$$\mathbf{P}_1 = \mathbf{P}_2 + t_1' \mathbf{b}_1 + \ldots + t_k' \mathbf{b}_k - s_1' \mathbf{a}_1 - \ldots - s_j' \mathbf{a}_j$$

$$= \mathbf{P}_2 + t_1' \mathbf{b}_1 + \ldots + t_k' \mathbf{b}_k - s_1' \sum_{m=1}^{k} \lambda_{1m} \mathbf{b}_m - \ldots - s_j' \sum_{m=1}^{k} \lambda_{jm} \mathbf{b}_m$$

12.9 $$= \mathbf{P}_2 + (t_1' - \sum_{i=1}^{j} s_i' \lambda_{i1}) \mathbf{b}_1 + \ldots + (t_k' - \sum_{i=1}^{j} s_i' \lambda_{ik}) \mathbf{b}_k$$

y por tanto $\mathbf{P}_1 \in \mathscr{P}_2$. Sea $\mathbf{P}$ un punto arbitrario en $\mathscr{P}_1$. Entonces para algunos $s_1, \ldots, s_j \in \mathrm{R}$, por 12.8 y 12.9, tenemos

$$\mathbf{P} = \mathbf{P}_1 + s_1 \mathbf{a}_1 + \ldots + s_j \mathbf{a}_j$$

$$= \mathbf{P}_1 + s_1 \sum_{m=1}^{k} \lambda_{1m} \mathbf{b}_m + \ldots + s_j \sum_{m=1}^{k} \lambda_{jm} \mathbf{b}_m$$

$$= \mathbf{P}_2 + [t_1' + \sum_{i=1}^{j} (s_i - s_i') \lambda_{i1}] \mathbf{b}_1 + \ldots + [t_k' + \sum_{i=1}^{j} (s_i - s_i') \lambda_{ik}] \mathbf{b}_k$$

luego $\mathbf{P} \in \mathscr{P}_2$. Hemos así demostrado que si $\mathscr{P}_1 \cap \mathscr{P}_2 \neq \varnothing$, entonces $\mathscr{P}_1 \subset \mathscr{P}_2$. Lo que completa la prueba.

12.10 Ejemplo. ¿Son paralelos los planos

$$\mathscr{P}_1 = \{(1, 4, -2, 3) + s_1(1, 0, 3, -12) + s_2(2, 5, -4, 6)\}$$

y

$$\mathscr{P}_2 = \{(2, 3, -2, 3) + t_1(1, -1, 3, -3) + t_2(0, -2, 2, 3) + t_3(0, -3, 4, -3)\}?$$

¿Es vacía su intersección?

SOLUCIÓN. Sean $\mathbf{a}_1 = (1, 0, 3, -12)$, $\mathbf{a}_2 = (2, 5, -4, 6)$, $\mathbf{b}_1 = (1, -1, 3, -3)$, $\mathbf{b}_2 = (0, -2, 2, 3)$, y $\mathbf{b}_3 = (0, -3, 4, -3)$. Los conjuntos $\{\mathbf{a}_1, \mathbf{a}_2\}$ y $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ son linealmente independientes. Los planos son paralelos si $\{\mathbf{a}_1, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ y $\{\mathbf{a}_2, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ son linealmente dependientes, es decir, si $\mathbf{a}_1$ y $\mathbf{a}_2$ son cada uno combinaciones lineales de $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$. Supongamos que $\mathbf{a}_1 = \lambda_{11}\mathbf{b}_1 + \lambda_{12}\mathbf{b}_2 + \lambda_{13}\mathbf{b}_3$, entonces

$$\begin{aligned} 1 &= \lambda_{11} \\ 0 &= -\lambda_{11} - 2\lambda_{12} - 3\lambda_{13} \\ 3 &= 3\lambda_{11} + 2\lambda_{12} + 4\lambda_{13} \\ -12 &= -3\lambda_{11} + 3\lambda_{12} - 3\lambda_{13}. \end{aligned}$$

Resolviendo estas ecuaciones encontramos $\lambda_{11} = 1$, $\lambda_{12} = -2$, $\lambda_{13} = 1$. Análogamente si $\mathbf{a}_2 = \lambda_{21}\mathbf{b}_1 + \lambda_{22}\mathbf{b}_2 + \lambda_{23}\mathbf{b}_3$, entonces

$$\begin{aligned} 2 &= \lambda_{21} \\ 5 &= -\lambda_{21} - 2\lambda_{22} - 3\lambda_{23} \\ -4 &= 3\lambda_{21} + 2\lambda_{22} + 4\lambda_{23} \\ 6 &= -3\lambda_{21} + 3\lambda_{22} - 3\lambda_{23}. \end{aligned}$$

Resolviendo estas ecuaciones, encontramos $\lambda_{21} = 2$, $\lambda_{22} = 1$, $\lambda_{23} = -3$. Como $\mathbf{a}_1$ y $\mathbf{a}_2$ son combinaciones lineales de $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$, los planos son paralelos.

Como, de acuerdo con el teorema 12.7, o $\mathscr{P}_1 \cap \mathscr{P}_2 = \mathscr{P}_1$ o $\mathscr{P}_1 \cap \mathscr{P}_2 = \varnothing$, es suficiente determinar si un punto cualquiera de $\mathscr{P}_1$ pertenece a $\mathscr{P}_2$. Sea $\mathbf{P}_1 = (1, 4, -2, 3) \in \mathscr{P}_1$. Si $\mathbf{P}_1 \in \mathscr{P}_2$ entonces hay números $t_1, t_2, t_3 \in \mathrm{R}$ tales que

$$(1, 4, -2, 3) = (2, 3, -2, 3) + t_1(1, -1, 3, -3) + t_2(0, -2, 2, 3) + t_3(0, -3, 4, -3).$$

Esta ecuación vectorial es equivalente al sistema de ecuaciones componentes

$$\begin{aligned} -1 &= t_1 \\ 1 &= -t_1 - 2t_2 - 3t_3 \\ 0 &= 3t_1 + 2t_2 + 4t_3 \\ 0 &= -3t_1 + 3t_2 - 3t_3. \end{aligned}$$

Resolviendo las primeras tres ecuaciones de este grupo encontramos que si el sistema tiene una solución ha de ser $t_1 = -1$, $t_2 = -\frac{9}{2}$, $t_3 = 3$. Sin embargo, estos valores no satisfacen la cuarta ecuación. Luego el sistema no tiene solución alguna. Por tanto $\mathbf{P}_1 \notin \mathscr{P}_2$ y, según el teorema 12.7, $\mathscr{P}_1 \cap \mathscr{P}_2 = \varnothing$.

Problemas

1. Determínese la dimensión del plano que pasa por el conjunto de puntos dado y determínese el plano

a) $\mathbf{P}_0 = (1, 1, 1, 1)$, $\mathbf{P}_1 = (1, 2, 3, 4)$
b) $\mathbf{P}_0 = (1, 1, 1, 1)$, $\mathbf{P}_1 = (0, -2, 2, -3)$, $\mathbf{P}_2 = (0, 3, 2, 1)$, $\mathbf{P}_3 = (1, 2, 3, 4)$
c) $\mathbf{P}_0 = (1, 0, 2, 0)$, $\mathbf{P}_1 = (0, 2, 3, 1)$, $\mathbf{P}_2 = (0, 0, 1, -1)$, $\mathbf{P}_3 = (1, 2, 4, 2)$
d) $\mathbf{P}_0 = (1, 0, 0, 0)$, $\mathbf{P}_1 = (0, 1, 0, 0)$, $\mathbf{P}_2 = (0, 0, 1, 2)$, $\mathbf{P}_3 = (0, 2, -1, -2)$

2. Encuéntrese la distancia entre los siguientes pares de puntos

a) $\mathbf{P}_0 = (1, 2, 3, 4)$, $\mathbf{P}_1 = (-1, 3, -2, 1)$
b) $\mathbf{P}_0 = (1, 1, 1, 1, 1)$, $\mathbf{P}_1 = (0, 2, -3, 7, -2)$
c) $\mathbf{P}_0 = (3, -2, 4, -6, 5)$, $\mathbf{P}_1 = (1, 0, 3, -2, 4)$
d) $\mathbf{P}_0 = (0, 0, 0, 0)$, $\mathbf{P}_1 = (1, 2, -2, -1)$.

3. Demuéstrese que los siguientes planos son paralelos y encuéntrese su intersección:

$$\mathscr{P}_1 = \{(-1, 2, 3, -2) + t(0, 1, 1, 1)\}$$
$$\mathscr{P}_2 = \{(1, 0, 2, 0) + t_1(1, -2, -1, -1) + t_2(1, 0, 1, 1)\}.$$

4. Demuéstrese que los siguientes planos son no paralelos y, sin embargo, tienen una intersección vacía. Explíquese por qué puede suceder esto.

$$\mathscr{P}_1 = \{(1, 1, 1, 1) + t(0, 1, 2, 3)\}$$
$$\mathscr{P}_2 = \{(1, 0, 2, 0) + t_1(1, -2, -1, -1) + t_2(1, 0, 1, 1)\}.$$

5. Demuéstrese que los siguientes planos tienen exactamente un punto en común:

$$\mathscr{P}_1 = \{(0, 0, 0, 0) + t_1(1, 0, 0, 0) + t_2(0, 1, 0, 0)\}$$
$$\mathscr{P}_2 = \{(0, 0, 0, 0) + t_3(0, 0, 1, 0) + t_4(0, 0, 0, 1)\}.$$

13. RESUMEN

En este capítulo comenzamos nuestro estudio de la geometría tridimensional. Estudiamos la geometría de las rectas y los planos. En capítulos posteriores estudiaremos la geometría de las curvas y de las superficies. Las curvas pueden considerarse como generalizaciones de las rectas y las superficies como generalizaciones de los planos.

El producto vectorial de dos vectores se definió en el espacio tridimensional. Se usó el producto vectorial y el triple producto escalar (con el producto vectorial íntimamente ligado) en la discusión de la independencia lineal de los vectores en el espacio tridimensional. Puede expresarse una normal a un plano como el producto vectorial de dos vectores linealmente independientes paralelos al plano, y la ecuación de un plano puede expresarse en términos de una normal al plano.

Después de una discusión de la intersección de planos y de la intersección de un plano con una recta, introdujimos el concepto de base de un espacio vectorial. A pesar de que confinamos nuestra discusión a bases en el espacio tridimensional V_3, el concepto puede extenderse a espacios vectoriales más generales. En la sección 12, dimos una breve introducción a los espacios euclidianos n-dimensionales.

Estos primeros dos capítulos complementan nuestro estudio del álgebra de los vectores y sirven como introducción a la rama del álgebra moderna llamada álgebra lineal. Aunque el material que aquí hemos presentado es suficiente para nuestros propósitos, nadie que se interese seriamente en estudios más avanzados de matemáticas, pura o aplicada, puede arreglárselas sin un conocimiento más extenso del álgebra lineal. Para facilitar tal extensión de conocimientos en esta rama fundamental de la matemática damos en la bibliografía de la página 777 una lista de algunos textos y tratados sobre ella. Lo que aquí se ha hecho esperamos prepare al lector para una apreciación adecuada de un desarrollo del tema más abstracto y sistemático. Pasamos ahora del álgebra vectorial a lo que podríamos llamar cálculo vectorial.

Problemas de repaso

1. Dados los puntos $\mathbf{P}_1 = (1, 1, 1)$, $\mathbf{P}_2 = (-2, 1, -3)$ y $\mathbf{P}_3 = (3, -2, 4)$:

a) encuéntrese la recta $\mathscr{L}_1$ que pasa por $\mathbf{P}_1$ y $\mathbf{P}_2$;

b) demuéstrese que los puntos $\mathbf{P}_1$, $\mathbf{P}_2$ y $\mathbf{P}_3$ no son colineales;

c) encuéntrese una ecuación del plano $\mathscr{P}$ que pasa por $\mathbf{P}_1$, $\mathbf{P}_2$ y $\mathbf{P}_3$;

d) encuéntrese el área del triángulo $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$;

e) encuéntrese la recta $\mathscr{L}_2$ que pasa por $\mathbf{P}_4 = (-5, 2, -1)$ y es ortogonal al plano $\mathscr{P}$ de la parte *c*;

f) encuéntrese la distancia del punto $\mathbf{P}_4 = (-5, 2, -1)$ al plano $\mathscr{P}$ de la parte *c*;

g) encuéntrese la intersección de la recta $\mathscr{L}_3$ que pasa por $\mathbf{P}_4 = (-5, 2, -1)$ y $\mathbf{P}_5 = (2, -1, 2)$ con el plano $\mathscr{P}$ de la parte *c*;

h) encuéntrese la recta $\mathscr{L}_4$ que pasa por $\mathbf{P}_4 = (-5, 2, -1)$ paralela a la recta $\mathscr{L}_1$ de la parte *a*;

i) encuéntrese un ángulo entre la recta $\mathscr{L}_2$ de la parte *e* y la recta $\mathscr{L}_3$ de la parte *g*;

j) determínense los cosenos directores de la recta $\mathscr{L}_2$ de la parte *e*.

Sugerencia. Véase la ecuación 3.12, pág. 53.

2. Dados los vectores $\mathbf{a} = (1, -2, 3)$, $\mathbf{b} = (1, 1, -2)$ y $\mathbf{c} = (1, -3, 4)$.

a) calcúlese el área del paralelogramo de lados **a** y **b**;
b) calcúlese el área del triángulo de lados **a**, **c** y $\mathbf{c}-\mathbf{a}$;
c) calcúlese el volumen del paralelepípedo de aristas **a**, **b** y **c**;
d) calcúlese el volumen del tetraedro con aristas **a**, **b** y **c**.

3. Establézcase la identidad

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}.$$

4. Si $|\mathbf{a}| = 3$ y $|\mathbf{b}| = 4$, evalúese

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \cdot \mathbf{b})^2.$$

5. Determínese un ángulo entre los planos de ecuaciones

$$3x+2y+5z = 3 \quad \text{y} \quad 6x+y+z = 7.$$





Funciones vectoriales de una variable real

1. INTRODUCCIÓN

Hemos definido la recta de R^3 que pasa por $\mathbf{P}_0$ y es paralela a $\mathbf{a}$ como el conjunto $\{\mathbf{P}_0 + t\mathbf{a} \mid t \in R\}$. En esta descripción de una recta a cada número real t corresponde el punto $\mathbf{P}_0 + t\mathbf{a}$ de R^3. Tal correspondencia se llama función vectorial de una variable real. Si denotamos a esta función por $\mathbf{f}$, entonces su regla de correspondencia es

$$\mathbf{f}(t) = \mathbf{P}_0 + t\mathbf{a} = (x_0 + ta_1, y_0 + ta_2, z_0 + ta_3)$$

donde $\mathbf{P}_0 = (x_0, y_0, z_0)$ y $\mathbf{a} = (a_1, a_2, a_3)$. El dominio de $\mathbf{f}$ es el conjunto de todos los números reales y el rango de $\mathbf{f}$ es la recta que pasa por $\mathbf{P}_0$ y es paralela a $\mathbf{a}$.

Cualquier función que tiene un conjunto de números reales como

dominio y un conjunto de vectores (o puntos) como su rango se llama función vectorial de una variable real. En este capítulo estudiaremos funciones de este tipo y consideraremos para tales funciones los conceptos de límite, continuidad, derivada e integral. Todo esto lleva consigo poco que sea realmente nuevo; en la mayor parte de los casos usamos las técnicas desarrolladas en el cálculo de funciones reales de una variable real.

Después de la discusión del cálculo de las funciones vectoriales damos algunas aplicaciones en geometría y física.

2. FUNCIONES VECTORIALES DE UNA VARIABLE REAL

Antes de discutir las funciones vectoriales, repasaremos brevemente la noción básica de función. Una *función* f es una correspondencia de un conjunto $\mathcal{A}$ en un conjunto $\mathcal{B}$ tal que a cada elemento de $\mathcal{A}$ corresponde un y sólo un elemento de $\mathcal{B}$. En las funciones descutidas en la primera parte del cálculo, $\mathcal{A}$ y $\mathcal{B}$ son conjuntos de números reales; tales funciones se llaman funciones reales de una variable real. Sin embargo, la noción básica de función no establece restricción alguna sobre el carácter de los objetos que comprenden los conjuntos $\mathcal{A}$ y $\mathcal{B}$. En este capítulo nos ocuparemos de las funciones en que $\mathcal{A}$ es un conjunto de números reales y $\mathcal{B}$ es un conjunto de vectores o puntos.

El conjunto $\mathcal{A}$ se llama **dominio** de la función y el conjunto de elementos de $\mathcal{B}$ que se corresponden con elementos de $\mathcal{A}$ se llama el rango de la función. Si la función se denota por f, entonces su dominio y su rango se denotan por $\mathcal{D}_f$ y $\mathcal{R}_f$ respectivamente. Además, $f(x)$ denota el elemento de $\mathcal{R}_f$ que corresponde al elemento x de $\mathcal{D}_f$; a $f(x)$ se le llama valor de f en x. En general, una función se especifica dando su dominio y una regla de correspondencia; es decir, una prescripción para determinar $f(x)$ para cada $x \in \mathcal{D}_f$.

Una *función vectorial de una variable real* es una función cuyo dominio es un conjunto de números reales y cuyo rango es un conjunto de vectores o puntos de R^n. Denotamos tales funciones por letras negrillas tales como **f**, **g**, etc.

Por ejemplo,

$$\mathbf{f}(t) = (1, 3, 2) + t(1, 1, 2) = (1+t, 3+t, 2+2t), \qquad \mathcal{D}_{\mathbf{f}} = R$$

describe una función vectorial de una variable real. El rango de esta función es una recta en R^3 y la función es una correspondencia o transformación de puntos sobre la recta real R en puntos sobre la recta que pasa por $(1, 3, 2)$ y es paralela a $(1, 1, 2)$. El punto 0 de R se transforma en $\mathbf{f}(0) = (1, 3, 2)$; -1 se transforma en $\mathbf{f}(-1) = (0, 2, 0)$; etc. Si escribimos $\mathbf{f}(t)$ en términos de sus componentes tenemos $\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t))$ donde $f_1(t) = 1+t, f_2(t) = 3+t, f_3(t) = 2+2t$. Las funciones f_1, f_2, f_3 se

llaman funciones componentes de la función **f**; estas funciones componentes son funciones reales de una variable real. Si I denota la función identidad en los números reales, $I(t) = t$, entonces $f_1 = 1+I$, $f_2 = 3+I$, y $f_3 = 2+2I$. Podemos pues escribir la función **f** en términos de componentes como sigue:

$$\mathbf{f} = (f_1, f_2, f_3) = (1+I, 3+I, 2+2I).$$

En general, si el rango de **f** es un conjunto de vectores en R^n, podemos escribir

$$\mathbf{f} = (f_1, f_2, \ldots, f_n)$$

donde $f_i(t)$ es el i-ésimo componente de $\mathbf{f}(t)$. La función real f_i con dominio $\mathscr{D}_\mathbf{f}$ se llama la i-ésima componente de la función vectorial **f**. De esta forma, una función vectorial **f** con rango en R^n define n funciones reales $f_1, f_2, \ldots, f_n$, todas las cuales tienen $\mathscr{D}_\mathbf{f}$ como dominio. Como veremos, esta representación de una función vectorial en términos de sus funciones componentes nos permite aplicar a las funciones vectoriales las técnicas desarrolladas en el cálculo de las funciones reales.

2.1 Ejemplo. Si $\mathbf{f} = (a \cosh, b \operatorname{senh})$ donde $a > 0$ y $b > 0$, demuéstrese que el rango de **f** es una rama de una hipérbola.

SOLUCIÓN. Un punto (x, y) pertenece al rango de **f** si y sólo si $x = a \cosh t$ y $y = b \operatorname{senh} t$ para algún $t \in R$. Así pues, si $(x, y) \in \mathscr{R}_\mathbf{f}$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \cosh^2 t - \operatorname{senh}^2 t = 1.$$

Esto nos demuestra que si $(x, y) \in \mathscr{R}_\mathbf{f}$ entonces (x, y) está sobre la hipérbola de ecuación $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$; en realidad, (x, y) está sobre la rama derecha de esta hipérbola ya que $x = a \cosh t > 0$. Llamemos a esta rama de la hipérbola $\mathscr{H}$. Ahora bien, si $(x, y) \in \mathscr{H}$, entonces existe un número t tal que $y = b \operatorname{senh} t$. Usando la ecuación para $\mathscr{H}$, obtenemos

$$\frac{x^2}{a^2} = 1 + \operatorname{senh}^2 t = \cosh^2 t.$$

Como x es positiva, concluimos que $x = a \cosh t$. Lo que nos demuestra que si $(x, y) \in \mathscr{H}$ entonces $(x, y) \in \mathscr{R}_\mathbf{f}$, y, por tanto, el rango de **f** es $\mathscr{H}$.

2.2 Ejemplo. Dibújese el rango de **f** cuando

$$\mathbf{f}(t) = (t, t, 2t^2) \quad \mathscr{D}_\mathbf{f} = [-3, 3].$$

SOLUCIÓN. El rango de **f** es el conjunto de puntos $\{\mathbf{f}(t) \mid \mathbf{f}(t) = (t, t, 2t^2), t \in [-3, 3]\}$. Si escribimos $\mathbf{f}(t) = t(1, 1, 0) + t^2(0, 0, 2)$, vemos que $\mathbf{f}(t)$ es la

suma de un vector a lo largo de la recta $y = x$ en el plano XY y un vector perpendicular al plano XY. Así pues el rango de **f** debe encontrarse en el plano de ecuación $y = x$. Podemos ver esto también del siguiente modo: Para cada punto (x, y, z) del rango de **f**, $x = t$, $y = t$, $z = 2t^2$. Como el plano con ecuación $y = x$ es el conjunto de todos los puntos (x, y, z) de R^3 tales que $y = x$, el rango de **f** debe encontrarse en tal plano.

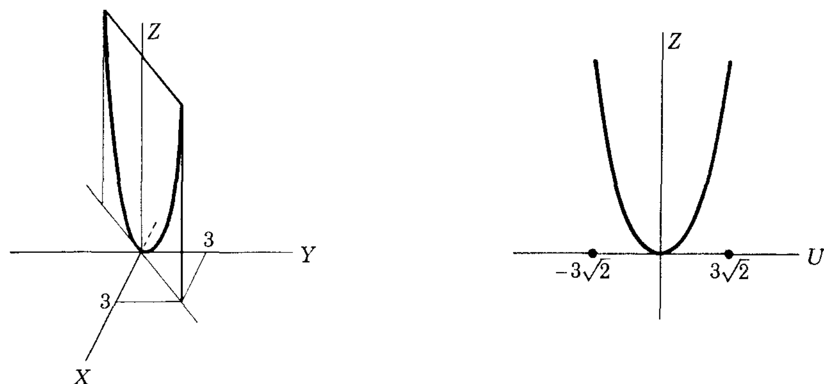


FIGURA 1

Si hacemos $u = x \sec \frac{\pi}{4} = \sqrt{\ }x$, entonces u es una distancia dirigida a lo largo de la recta $y = x$ en el plano XY. El rango de **f** es una porción de la parábola $z = u^2$ que se encuentra en el plano que contiene el eje Z y la recta $y = x$ en el plano XY (figura 1).

Problemas

1. Proporciónese una función del intervalo $[0, 1]$ sobre el segmento rectilíneo que une los puntos

a) $(-1, 2)$ y $(3, 5)$ *b*) $\mathbf{P}_0$ y $\mathbf{P}_1$ en R^2
c) $(1, 4, 7)$ y $(3, -2, 1)$ *c*) $\mathbf{P}_0$ y $\mathbf{P}_1$ en R^n.

2. *a*) Proporciónese una función del intervalo $[-2, 3]$ sobre el segmento rectilíneo que une los puntos $(1, 0, 3)$ y $(4, 2, -1)$.

b) Proporciónese una función del intervalo $[a, b]$ sobre el segmento rectilíneo $[\mathbf{P}_0, \mathbf{P}_1]$.

3. Si $\mathbf{f}(t) = (a \cos t, a \operatorname{sen} t)$ donde $a > 0$ y $\mathscr{D}_{\mathbf{f}} = [0, 2\pi]$, demuéstrese que el rango de **f** es una circunferencia en R^2.

4. Si $\mathbf{f}(t) = (a \cos t, b \operatorname{sen} t)$ donde $a > 0$, $b > 0$, y $\mathscr{D}_{\mathbf{f}} = [0, 2\pi]$, demuéstrese que el rango de **f** es una elipse en R^2. Si $a = 2$ y $b = 4$, dibújese la elipse.

5. Si $\mathbf{f} = (3I, I^2)$, demuéstrese que el rango de **f** es una parábola en R^2.

6. Si $\mathbf{f} = \left(\dfrac{1-I}{1+I^2}, \dfrac{2I}{1+I^2}\right)$ descríbase el rango de **f**.

7. Dibújese el rango de **f** cuando $\mathbf{f}(t) = (t, t, \operatorname{sen} t)$, $t \in [0, 4\pi]$.

3. EL LÍMITE DE UNA FUNCIÓN VECTORIAL

En esta sección extenderemos el concepto de límite de una función real de una variable real a las funciones vectoriales de una variable real. Pero revisemos primero la definición de límite de una función real.

Sea f una función real de una variable real y sea a un punto de acumulación de $\mathscr{D}_f$ (el punto a es un punto de acumulación de $\mathscr{D}_f$ si todo intervalo abierto que contiene a contiene un punto t en $\mathscr{D}_f$ distinto de a). Se dice que un número b es el *límite de la función f en a* si para cada número $\varepsilon > 0$, hay un número $\delta > 0$ tal que siempre que $t \in \mathscr{D}_f$ y $0 < |t-a| < \delta$ entonces

$$|f(t)-b| < \varepsilon.$$

Si b es el límite de f en a, escribimos

$$\lim_a f = b \quad \text{o} \quad \lim_{t\to a} f(t) = b.$$

El término $|t-a|$ es la distancia de t a a y $|f(t)-b|$ es la distancia de $f(t)$ a b sobre la recta real R. Así pues, la definición de $\lim_a f = b$ afirma que $f(t)$ permanece arbitrariamente próximo a b para todo t suficientemente próximo a a, pero distinto de a. Para funciones vectoriales el concepto de límite tiene el mismo significado intuitivo: $\lim_a \mathbf{f} = \mathbf{b}$ si $\mathbf{f}(t)$ permanece arbitrariamente próximo a **b** para t suficientemente próximo a a, pero distinto de a. Como en R^n la distancia de $\mathbf{f}(t)$ a **b** es $|\mathbf{f}(t)-\mathbf{b}|$, la definición de $\lim_a \mathbf{f} = \mathbf{b}$ es por analogía:

3.1 Definición. *Se dice que el vector* **b** *es el* ***límite de la función* f *en*** a *si para cada número $\varepsilon > 0$ existe un número $\delta > 0$ tal que siempre que t está en el dominio de* **f** *y $0 < |t-a| < \delta$ entonces*

$$|\mathbf{f}(t)-\mathbf{b}| < \varepsilon.$$

Las notaciones

$$\lim_a \mathbf{f} = \mathbf{b} \quad \text{y} \quad \lim_{t\to a} \mathbf{f}(t) = \mathbf{b}$$

se usan para denotar que **b** es el límite de **f** en a. Siempre que se trate del límite de **f** en a se está suponiendo que a es un punto de acumulación de $\mathscr{D}_\mathbf{f}$.

Nota. Obsérvese que el límite vectorial $\lim_{t \to a} \mathbf{f}(t) = \mathbf{b}$ es equivalente al límite real $\lim_{t \to a} |\mathbf{f}(t) - \mathbf{b}| = 0$; es decir, cuando t tiende a a, $\mathbf{f}(t)$ tiende a $\mathbf{b}$ si y sólo si la longitud del vector $\mathbf{f}(t) - \mathbf{b}$ tiende a 0.

Para dar un sentido más geométrico a la definición de límite, introducimos la noción de vecindad en R_n. Una *vecindad* de $\mathbf{c}$ de radio r es el interior de la esfera n-dimensional de radio r y centro $\mathbf{c}$: $\mathscr{S}(\mathbf{c}; r) = \{x \mid |x - c| < r\}$. Así pues, en R una vecindad es un intervalo abierto, es decir

$$\mathscr{S}(c; r) = \{x \mid |x - c| < r\} = \langle c - r, c + r \rangle.$$

En R^2, una vecindad es el interior de un círculo y en R^3 es el interior de una esfera. Si omitimos el punto $\mathbf{c}$ de la vecindad $\mathscr{S}(\mathbf{c}; r)$, entonces tenemos una *vecindad reducida* de $\mathbf{c}$; a esta vecindad reducida la denotamos por $\mathscr{S}'(\mathbf{c}; r)$.

En términos de vecindades la definición de $\lim_{a} \mathbf{f} = \mathbf{b}$ dice: $\mathbf{b}$ es el límite de $\mathbf{f}$ en a si para cada vecindad $\mathscr{S}(\mathbf{b}; \varepsilon)$ de $\mathbf{b}$, existe una vecindad reducida $\mathscr{S}'(a; \delta)$ de a tal que $\mathbf{f}$ transforma $\mathscr{S}'(a; \delta)$ en $\mathscr{S}(\mathbf{b}; \varepsilon)$.[1]

3.2 Ejemplo. Si $\mathbf{f} = (3I, I^2)$, determínese $\lim_{2} \mathbf{f}$.

Solución. Para t próximo a 2, vemos que $\mathbf{f}(t) = (3t, t^2)$ está próximo a (6,4). Así pues, suponemos que $\lim_{2} \mathbf{f} = (6, 4)$ y verificamos después que tal es el caso. Sea $\varepsilon > 0$ cualquiera. Queremos encontrar una $\delta > 0$ tal que $|(3t, t^2) - (6, 4)| < \varepsilon$ siempre que $0 < |t - 2| < \delta$. Ahora bien

$$|(3t, t^2) - (6, 4)| = [(3t - 6)^2 + (t^2 - 4)^2]^{1/2};$$

por tanto

$$|(3t, t^2) - (6, 4)| < \varepsilon \quad \text{si} \quad |3t - 6| < \frac{\varepsilon}{\sqrt{2}} \text{ y } |t^2 - 4| < \frac{\varepsilon}{\sqrt{2}}.$$

Como el $\lim_{t \to 2} 3t = 6$, existe una $\delta_1 > 0$ $\left(\text{por ejemplo, } \delta_1 = \frac{\varepsilon}{3\sqrt{2}}\right)$ tal que $|3t - 6| < \frac{\varepsilon}{\sqrt{2}}$ siempre que $0 < |t - 2| < \delta_1$. Por otra parte, $\lim_{t \to 2} t^2 = 4$ implica que existe una $\delta_2 > 0$ $\left(\text{por ejemplo, } \delta_2 = \text{mín}\left\{1, \frac{\varepsilon}{5\sqrt{2}}\right\}\right)$ tal que $|t^2 - 4| < \frac{\varepsilon}{\sqrt{2}}$ siempre que $0 < |t - 2| < \delta_2$. Luego si $\delta = \text{mín}\{\delta_1, \delta_2\}$,

$$|(3t, t^2) - (6, 4)| = [(3t - 6)^2 + (t^2 - 4)^2]^{1/2} < \left[\left(\frac{\varepsilon}{\sqrt{2}}\right)^2 + \left(\frac{\varepsilon}{\sqrt{2}}\right)^2\right]^{1/2} = \varepsilon$$

[1] Es claro que lo que realmente transforma $\mathbf{f}$ en $\mathscr{S}(\mathbf{b}; \varepsilon)$ es $\mathscr{S}'(a; \delta) \cap \mathscr{D}_\mathbf{f}$, conjunto nunca vacío por ser a de acumulación de $\mathscr{D}_\mathbf{f}$.[N. de/T].

siempre que $0 < |t-2| < \delta$. Con lo que hemos verificado que $\lim_{t\to 2} (3t, t^2) = (6, 4)$.

Utilizamos la figura 2 para dar ahora una interpretación geométrica de la solución del ejemplo 3.2. Escogemos una δ tal que siempre que t se encuentre a una distancia menor que δ de 2, las longitudes de los lados del rectángulo sean menores que $\varepsilon/\sqrt{2}$. Entonces la longitud de la diagonal debe ser menor que ε.

En el ejemplo 3.2 el límite de la función vectorial es el vector cuyos componentes son los límites de los correspondientes componentes de la función. Esto es cierto para cualquier función vectorial y la prueba de este hecho es esencialmente la misma que la prueba dada en la solución del ejemplo 3.2 a que $\lim_{t\to 2} (3t, t^2) = (6, 4)$.

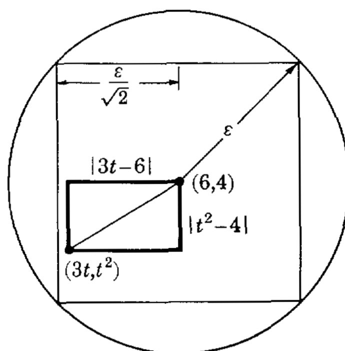


FIGURA 2

3.3 Teorema. *Sea* $\mathbf{b} = (b_1, \ldots, b_n) \in \mathrm{R}^n$, $\mathbf{f} = (f_1, \ldots, f_n)$ *una función de* R *en* R^n, *y* a *un punto de acumulación de* $\mathscr{D}_{\mathbf{f}}$. *Entonces* $\lim_a \mathbf{f} = \mathbf{b}$ *si y sólo si* $\lim_a f_i = b_i$ *para* $i = 1, \ldots, n$.

PRUEBA. Si $\lim_a \mathbf{f} = \mathbf{b}$, entonces para cualquier $\varepsilon > 0$ existe una $\delta > 0$ tal que

$$|\mathbf{f}(t) - \mathbf{b}| = \left[\sum_{i=1}^{n} (f_i(t) - b_i)^2\right]^{1/2} < \varepsilon$$

siempre que $t \in \mathscr{D}_{\mathbf{f}}$ y $0 < |t-a| < \delta$. De donde

$$|f_i(t) - b_i| < \varepsilon \text{ para cada } i = 1, \ldots, n$$

siempre que $t \in \mathscr{D}_{\mathbf{f}} = \mathscr{D}_{f_i}$ y $0 < |t-a| < \delta$. Esto nos muestra que $\lim \mathbf{f} = \mathbf{b}$ implica que $\lim_a f_i = b_i$ $(i = 1, \ldots, n)$.

Si $\lim_a f_i = b_i$ para cada $i = 1, \ldots, n$, entonces para cualquier $\varepsilon > 0$ existe una $\delta_i > 0$ tal que

$$|f_i(t) - b_i| < \frac{\varepsilon}{\sqrt{n}}$$

siempre que $t \in \mathscr{D}_{f_i}$ y $0 < |t-a| < \delta_i$. Tomando $\delta = \text{mín}\,\{\delta_1, \ldots, \delta_n\}$, tenemos que

$$|\mathbf{f}(t) - \mathbf{b}| = \left[\sum_{i=1}^{n} (f_i(t) - b_i)^2\right]^{1/2} < \left[\sum_{i=1}^{n} \left(\frac{\varepsilon}{\sqrt{n}}\right)^2\right]^{1/2} = \varepsilon$$

siempre que $t \in \mathscr{D}_\mathbf{f}$ y $0 < |t-a| < \delta$; es decir, que $\lim_a \mathbf{f} = \mathbf{b}$. Y esto completa la prueba.

El teorema 3.3 nos dice que (si el límite existe):

$$\lim_a \mathbf{f} = (\lim_a f_1, \ldots, \lim_a f_n).$$

El límite de una función vectorial $\mathbf{f}$ puede, por tanto, calcularse de acuerdo con los límites de las funciones reales componentes f_i. Por ejemplo,

$$\lim_{t \to \frac{\pi}{4}} (t, \text{sen}\, t, \tan t) = \left(\lim_{t \to \frac{\pi}{4}} t, \lim_{t \to \frac{\pi}{4}} \text{sen}\, t, \lim_{t \to \frac{\pi}{4}} \tan t\right)$$

$$= \left(\frac{\pi}{4}, \frac{\sqrt{2}}{2}, 1\right).$$

El teorema 3.3 nos permite probar algunos teoremas sobre límites de funciones vectoriales usando teoremas muy conocidos sobre límites de funciones reales; por ejemplo, el límite de una suma es la suma de los límites (si los límites existen). Pero antes de presentar estos teoremas definiremos algunas operaciones sobre funciones vectoriales.

3.4 Definición. *Si* $\mathbf{f}$ *y* $\mathbf{g}$ *son funciones vectoriales con rangos en* R^n *y dominios* $\mathscr{D}_\mathbf{f}$ *y* $\mathscr{D}_\mathbf{g}$ *en* R, *entonces* $\mathbf{f}+\mathbf{g}$, $\mathbf{f}-\mathbf{g}$, $\mathbf{f} \cdot \mathbf{g}$, *y* $\mathbf{f} \times \mathbf{g}$ *son funciones con dominio* $\mathscr{D}_\mathbf{f} \cap \mathscr{D}_\mathbf{g}$ *y reglas de correspondencia:*

$$[\mathbf{f}+\mathbf{g}](t) = \mathbf{f}(t) + \mathbf{g}(t)$$
$$[\mathbf{f}-\mathbf{g}](t) = \mathbf{f}(t) - \mathbf{g}(t)$$
$$[\mathbf{f}\cdot\mathbf{g}](t) = \mathbf{f}(t) \cdot \mathbf{g}(t)$$
$$[\mathbf{f}\times\mathbf{g}](t) = \mathbf{f}(t) \times \mathbf{g}(t) \quad \textit{(definida solamente si } \mathrm{R}^n = \mathrm{R}^3).$$

Si $\mathbf{f}$ *es una función vectorial y* φ *es una función real de una variable real, entonces la función* $\varphi\mathbf{f}$ *está definida como sigue:*

$$[\varphi\mathbf{f}](t) = \varphi(t)\,\mathbf{f}(t) \quad \mathscr{D}_{\varphi\mathbf{f}} = \mathscr{D}_\varphi \cap \mathscr{D}_\mathbf{f}.$$

Estas operaciones pueden también expresarse en términos de las funciones componentes. Si $\mathbf{f} = (f_1, \ldots, f_n)$ y $\mathbf{g} = (g_1, \ldots, g_n)$ entonces para cualquier $t \in \mathscr{D}_\mathbf{f} \cap \mathscr{D}_\mathbf{g}$

$$\begin{aligned}[\mathbf{f}+\mathbf{g}](t) &= \mathbf{f}(t)+\mathbf{g}(t) \\ &= (f_1(t), \ldots, f_n(t))+(g_1(t), \ldots, g_n(t)) \\ &= (f_1(t)+g_1(t), \ldots, f_n(t)+g_n(t)) \\ &= ([f_1+g_1](t), \ldots, [f_n+g_n](t)).\end{aligned}$$

Por tanto

3.5 $$\mathbf{f}+\mathbf{g} = (f_1+g_1, \ldots, f_n+g_n).$$

Análogamente podemos mostrar que

3.6 $$\mathbf{f}-\mathbf{g} = (f_1-g_1, \ldots, f_n-g_n)$$

3.7 $$\mathbf{f}\cdot\mathbf{g} = \sum_{i=1}^{n} f_i g_i$$

3.8 $$\varphi\mathbf{f} = (\varphi f_1, \ldots, \varphi f_n).$$

Si $\mathbf{f} = (f_1, f_2, f_3)$ y $\mathbf{g} = (g_1, g_2, g_3)$, entonces

3.9 $$\mathbf{f}\times\mathbf{g} = (f_2g_3-f_3g_2, f_3g_1-f_1g_3, f_1g_2-f_2g_1).$$

Nótese que $\mathbf{f}\cdot\mathbf{g}$ es una función real de variable real. Por ejemplo, si $\mathbf{f} = (I, \cos, \text{sen})$ y $\mathbf{g} = (\exp, I^{1/2}, I^2)$, entonces

$$\mathbf{f}\cdot\mathbf{g} = I\exp + I^{1/2}\cos + I^2\,\text{sen},$$

es decir,

$$[\mathbf{f}\cdot\mathbf{g}](t) = te^t+\sqrt{t}\cos t+t^2\,\text{sen}\, t.$$

Como $\mathscr{D}_\mathbf{f} = \mathrm{R}$ y $\mathscr{D}_\mathbf{g} = [0, \infty\rangle$ $(\mathscr{D}_{I^{\frac{1}{2}}} = [0, \infty\rangle)$, entonces

$$\mathscr{D}_{\mathbf{f}\cdot\mathbf{g}} = \mathscr{D}_\mathbf{f} \cap \mathscr{D}_\mathbf{g} = [0, \infty\rangle.$$

Usando el teorema 3.3, podemos probar fácilmente los siguientes teoremas.

3.10 Teorema. *Si* $\mathbf{f}$ *y* $\mathbf{g}$ *son funciones vectoriales de una variable real tales que*

$$\lim_a \mathbf{f} = \mathbf{b} \quad y \quad \lim_a \mathbf{g} = \mathbf{c}$$

y a es un punto de acumulación de $\mathscr{D}_\mathbf{f} \cap \mathscr{D}_\mathbf{g}$,[1] *entonces*

$$\lim_a [\mathbf{f}+\mathbf{g}] = \lim_a \mathbf{f} + \lim_a \mathbf{g} = \mathbf{b}+\mathbf{c}$$

[1] Obsérvese que esta condición es necesaria. Podría ser, en efecto, que a fuese de acumulación de $\mathscr{D}_\mathbf{f}$ y de $\mathscr{D}_\mathbf{g}$, pero no de $\mathscr{D}_\mathbf{f} \cap \mathscr{D}_\mathbf{g}$. [N. del T.]

$$\lim_a [\mathbf{f}-\mathbf{g}] = \lim_a \mathbf{f} - \lim_a \mathbf{g} = \mathbf{b}-\mathbf{c}$$

$$\lim_a [\mathbf{f}\cdot\mathbf{g}] = \left(\lim_a \mathbf{f}\right)\cdot\left(\lim_a \mathbf{g}\right) = \mathbf{b}\cdot\mathbf{c}$$

$$\lim_a [\mathbf{f}\times\mathbf{g}] = \left(\lim_a \mathbf{f}\right)\times\left(\lim_a \mathbf{g}\right) = \mathbf{b}\times\mathbf{c} \quad \textit{(para } \mathrm{R}^3 \textit{ solamente)}.$$

PRUEBA. Solamente probaremos la fórmula de la suma. Las pruebas de las otras partes son análogas. Si $\mathbf{f} = (f_1, \ldots, f_n)$ y $\mathbf{g} = (g_1, \ldots, g_n)$,

$$\begin{aligned}
\lim_a [\mathbf{f}+\mathbf{g}] &= \lim_a (f_1+g_1, \ldots, f_n+g_n) \\
&= \left(\lim_a (f_1+g_1), \ldots, \lim_a (f_n+g_n)\right) \\
&= \left(\lim_a f_1 + \lim_a g_1, \ldots, \lim_a f_n + \lim_a g_n\right) \\
&= \left(\lim_a f_1, \ldots, \lim_a f_n\right) + \left(\lim_a g_1, \ldots, \lim_a g_n\right) \\
&= \lim_a (f_1, \ldots, f_n) + \lim_a (g_1, \ldots, g_n) \\
&= \lim_a \mathbf{f} + \lim_a \mathbf{g}.
\end{aligned}$$

3.11 Teorema. *Si* $\mathbf{f}$ *es una función vectorial y* φ *es una función real y*

$$\lim_a \mathbf{f} = \mathbf{b} \quad y \quad \lim_a \varphi = r$$

a es un punto de acumulación de $\mathscr{D}_{\varphi\mathbf{f}}$, *entonces*

$$\lim_a (\varphi\mathbf{f}) = \lim_a \varphi \lim_a \mathbf{f} = r\mathbf{b}.$$

PRUEBA. Si $\mathbf{f} = (f_1, \ldots, f_n)$, *entonces*

$$\begin{aligned}
\lim_a (\varphi\mathbf{f}) &= \lim_a (\varphi f_1, \ldots, \varphi f_n) \\
&= \left(\lim_a (\varphi f_1), \ldots, \lim_a (\varphi f_n)\right) \\
&= \left(\lim_a \varphi \lim_a f_1, \ldots, \lim_a \varphi \lim_a f_n\right) \\
&= \lim_a \varphi \left(\lim_a f_1, \ldots, \lim_a f_n\right) \\
&= \lim_a \varphi \lim_a \mathbf{f}.
\end{aligned}$$

El teorema 3.10 nos dice que para las funciones vectoriales el límite de la suma es la suma de los límites, el límite de la diferencia es la diferencia

de los límites, el límite del producto escalar es el producto escalar de los límites, y el límite del producto vectorial es el producto vectorial de los límites, con tal de que, en todos los casos, existan los límites de las funciones. El teorema 3.11 afirma que el límite de una función real por una función vectorial es el límite de la función real por el límite de la función vectorial, si es que los límites de estas funciones existen.

Nota. Los teoremas 3.10 y 3.11 podrían haberse probado directamente partiendo de la definición de límite. Las pruebas habrían sido análogas a las de los correspondientes teoremas para funciones reales, ya que la longitud de un vector tiene las mismas propiedades básicas que el valor absoluto de un número real (problema 7).

Problemas

1. Si $\mathbf{f}(t) = (t, t^2)$, calcúlese y márquese la posición de $\mathbf{f}(0.9)$, $\mathbf{f}(0.99)$, $\mathbf{f}(0.999)$, $\mathbf{f}(1.1)$, $\mathbf{f}(1.01)$, $\mathbf{f}(1.001)$. Úsese la definición 3.1 para verificar que $\lim_{t\to 1} \mathbf{f}(t) = (1, 1)$.

2. Si $\mathbf{f}(t) = ([t], t)$, calcúlense y márquense las posiciones de $\mathbf{f}(1.9)$, $\mathbf{f}(1.99)$, $\mathbf{f}(2.1)$, $\mathbf{f}(2.01)$. ($[t]$ es el mayor entero no mayor que t.) Demuéstrese que $\lim_{t\to 2} \mathbf{f}(t)$ no existe.

3. Determínese $\lim_a \mathbf{f}$ (si es que existe), cuando

a) $\mathbf{f} = (I^{1/2}, I^2, \text{sen})$, $a = 2$

b) $\mathbf{f} = (\exp, \text{senh}, \cosh)$, $a = 1$

c) $\mathbf{f}(t) = \left(\ln t, \sqrt{1+t^2}, \dfrac{2t}{4-t^2}\right)$, $a = 2$

d) $\mathbf{f}(t) = \left(\dfrac{t}{1+t^2}, \dfrac{1+2t}{t^2}, 3t^2\right)$, $a = 3$.

4. Si $\mathbf{f}(t) = ([t], t)$, determínese $\lim_{t\to 2^-} \mathbf{f}(t)$ y $\lim_{t\to 2^+} \mathbf{f}(t)$.

El límite a la izquierda de $\mathbf{f}$ *en* a *es* $\mathbf{b}$, *lo que se escribe* $\lim_{t\to a^-} \mathbf{f}(t) = \mathbf{b}$, *si para todo* $\varepsilon > 0$ *existe un* $\delta > 0$ *tal que* $|\mathbf{f}(t)-\mathbf{b}| < \varepsilon$ *siempre que* $t \in \mathscr{D}_{\mathbf{f}} \cap \langle a-\delta, a\rangle$; la definición de $\lim_{t\to a^+} \mathbf{f}(t) = \mathbf{b}$ llamado el límite a la derecha de $\mathbf{f}$ en a, es análoga.

5. Si a es un punto de acumulación de $\mathscr{D}_{\mathbf{f}}$ y existe $\lim_a \mathbf{f}$ demuéstrese que este límite es único; es decir, demuéstrese que si $\lim_a \mathbf{f} = \mathbf{b}$ y $\lim_a \mathbf{f} = \mathbf{c}$, entonces $\mathbf{b} = \mathbf{c}$.

6. *a*) Si $\lim_{t\to a} \mathbf{f} = \mathbf{b}$, demuéstrese que la longitud de $\mathbf{f}(t)$ se aproxima a la longitud de $\mathbf{b}$ a medida que t se aproxima a a, es decir, que $\lim_{t\to a} |\mathbf{f}(t)| = |\mathbf{b}|$.

b) Si $\lim_{t\to a} \mathbf{f}(t) = \mathbf{b} \neq \mathbf{0}$, demuéstrese que la dirección de $\mathbf{f}(t)$ se aproxima a la dirección de $\mathbf{b}$ a medida que t se aproxima a a, es decir, que

$$\lim_{t\to a} \frac{\mathbf{f}(t)}{|\mathbf{f}(t)|} = \frac{\mathbf{b}}{|\mathbf{b}|}.$$

7. Pruébese directamente partiendo de la definición de límite 3.1 que si a es un punto de acumulación $\mathscr{D}_{\mathbf{f}+\mathbf{g}}$ y $\lim_a \mathbf{f} = \mathbf{b}$ y $\lim_a \mathbf{g} = \mathbf{c}$, entonces $\lim_a (\mathbf{f}+\mathbf{g}) = \mathbf{b}+\mathbf{c}$.

8. Úsese el teorema 3.3 para probar que si a es un punto de acumulación de $\mathscr{D}_{\mathbf{f}\cdot\mathbf{g}}$ y $\lim_a \mathbf{f} = \mathbf{b}$ y $\lim_a \mathbf{g} = \mathbf{c}$, entonces $\lim_a (\mathbf{f}\cdot\mathbf{g}) = \mathbf{b}\cdot\mathbf{c}$.

9. Si $\mathbf{f}(t) = (t, t^2, t^3)$, determínese $\lim_{h\to 0} \dfrac{\mathbf{f}(t+h)-\mathbf{f}(t)}{h}$.

4. CONTINUIDAD

La extensión de la noción de continuidad del caso de funciones reales al de funciones vectoriales es tan natural y directa como la extensión del concepto de límite.

4.1 Definición. *La función* $\mathbf{f}$ *es* ***continua en el punto*** a *de* $\mathscr{D}_\mathbf{f}$ *si para cada* $\varepsilon > 0$ *existe una* $\delta > 0$ *tal que*

$$|\mathbf{f}(t)-\mathbf{f}(a)| < \varepsilon$$

siempre que $t\in\mathscr{D}_\mathbf{f}$ *y* $|t-a| < \delta$.

Si a no es un punto de acumulación de $\mathscr{D}_\mathbf{f}$, entonces $\mathbf{f}$ es continua en a, pues en este caso hay una $\delta > 0$ tal que a es el único punto en $\mathscr{D}_\mathbf{f} \cap \langle a-\delta, a+\delta\rangle$, y entonces, para cualquier $\varepsilon > 0$, $|\mathbf{f}(t)-\mathbf{f}(a)| < \varepsilon$ siempre que $t\in\mathscr{D}_\mathbf{f} \cap \langle a-\delta, a+\delta\rangle$.

Si a es un punto de acumulación de $\mathscr{D}_\mathbf{f}$, entonces la definición 4.1 es equivalente a: *la función* $\mathbf{f}$ *es continua en el punto* a *de* $\mathscr{D}_\mathbf{f}$ *si*

$$\lim_a \mathbf{f} = \mathbf{f}(a).$$

El siguiente teorema es una consecuencia inmediata del teorema 3.3 (pág. 103)

4.2 Teorema. *Si* $\mathbf{f} = (f_1, \ldots, f_n)$ y $a \in \mathscr{D}_{\mathbf{f}}$, *entonces* $\mathbf{f}$ *es continua en a si y sólo si* f_i *es continua en a, para todo* $i = 1, \ldots, n$.

PRUEBA. Si a no es un punto de acumulación de $\mathscr{D}_{\mathbf{f}}$, entonces la prueba es inmediata (recuérdese que para todo i, $\mathscr{D}_{f_i} = \mathscr{D}_{\mathbf{f}}$). Supongamos que a es un punto de acumulación de $\mathscr{D}_{\mathbf{f}}$. Según el teorema 3.3, $\lim_a \mathbf{f} = \mathbf{f}(a)$ si y sólo si $\lim_a f_i = f_i(a)$ para todo $i = 1, \ldots, n$. Lo que completa la prueba.

Así pues, la continuidad de una función vectorial en un punto a puede determinarse comprobando la continuidad de las funciones componentes en a. Por ejemplo, la función $\mathbf{f} = (I, \cos, \operatorname{sen})$ es continua en todos los puntos de R ya que I, cos y sen son continuas en todos los puntos de R.

Correspondiéndose con los teoremas 3.10 y 3.11 sobre límites, tenemos el siguiente teorema sobre continuidad.

4.3 Teorema. *Si las funciones* $\mathbf{f}$ *y* $\mathbf{g}$ *son continuas en a, entonces* $\mathbf{f}+\mathbf{g}$, $\mathbf{f}-\mathbf{g}$, $\mathbf{f} \cdot \mathbf{g}$ *y* $\mathbf{f} \times \mathbf{g}$ *son continuas en a. Si* $\mathbf{f}$ *y* φ *son continuas en a, entonces* $\varphi\mathbf{f}$ *es continua en a.*

PRUEBA. Probaremos solamente que $\mathbf{f}+\mathbf{g}$ es continua en a. Las pruebas para las restantes operaciones son análogas. Si a no es un punto de acumulación de $\mathscr{D}_{\mathbf{f}+\mathbf{g}}$, entonces $\mathbf{f}+\mathbf{g}$ es continua en a. Si a es un punto de acumulación de $\mathscr{D}_{\mathbf{f}+\mathbf{g}}$, entonces a es un punto de acumulación de $\mathscr{D}_{\mathbf{f}}$ y de $\mathscr{D}_{\mathbf{g}}$ y $\lim_a \mathbf{f} = \mathbf{f}(a)$ y $\lim_a \mathbf{g} = \mathbf{g}(a)$. Y tenemos entonces, de acuerdo con el teorema 3.10,

$$\lim_a [\mathbf{f}+\mathbf{g}] = \lim_a \mathbf{f} + \lim_a \mathbf{g} = \mathbf{f}(a)+\mathbf{g}(a) = [\mathbf{f}+\mathbf{g}](a).$$

Luego $\mathbf{f}+\mathbf{g}$ es eontinua en a.

4.4 Definición. *La función* $\mathbf{f}$ *es* ***continua sobre un conjunto*** $\mathscr{S} \subset \mathscr{D}_{\mathbf{f}}$ *si la función restringida* $\mathbf{f}_{\mathscr{S}}$ *es continua en cada uno de los puntos de* $\mathscr{S}$.

Como función restringida, $\mathbf{f}_{\mathscr{S}}$, donde $\mathscr{S} \subset \mathscr{D}_{\mathbf{f}}$, entendemos la función con dominio $\mathscr{S}$ y regla de correspondencia $\mathbf{f}_{\mathscr{S}}(t) = \mathbf{f}(t)$ para todo $t \in \mathscr{S}$.

En la mayoría de los casos de interés el conjunto $\mathscr{S}$ es un intervalo. Si $\mathscr{S}$ es un intervalo abierto, entonces la definición 4.4 es equivalente a: *la functión* $\mathbf{f}$ *es continua sobre el intervalo abierto* $\mathscr{I}$ *si* $\mathbf{f}$ *es continua en cada punto de* $\mathscr{I}$. Si $\mathscr{S}$ es un intervalo cerrado, entonces la definición 4.4 es equivalente a: *la función* $\mathbf{f}$ *es continua sobre el intervalo cerrado* $[a, b]$ *si* $\mathbf{f}$ *es continua sobre el intervalo abierto* $\langle a, b \rangle$ *y si* $\lim_{a^+} \mathbf{f} = \mathbf{f}(a)$ *y* $\lim_{b^-} \mathbf{f} = \mathbf{f}(b)$.

Una función se llama *continua* si es continua en cada punto de su dominio.

Problemas

1. Encuéntrense los puntos (si es que hay algunos) donde las siguientes funciones no son continuas y delinéese el rango de cada función.

a) $\mathbf{f} = (\exp, I)$, $\mathscr{D}_{\mathbf{f}} = [0, 2]$

b) $$\begin{cases} \mathbf{f}(t) = \left(t, \dfrac{\operatorname{sen} t}{t}\right), & t \in \langle 0, \pi] \\ \mathbf{f}(0) = (0, 1) \end{cases}$$

c) $\mathbf{f}(t) = (t, t, [t])$, $t \in [0, 4]$.

2. Si $\mathbf{f}(t) = (|t|, 2|t|, t)$, $t \in [-2, 2]$

y

$$\mathbf{g}(t) = \begin{cases} (-t, -2t, t), & t \in [-2, 0] \\ (2-t, 4-2t, 2-t), & t \in \langle 0, 2], \end{cases}$$

delinéese el rango de **f** y **g**.

3. Si **f** es continua sobre un conjunto $\mathscr{S}$ de muéstrese que $|\mathbf{f}|$ es continua sobre $\mathscr{S}$. La función $|\mathbf{f}|$ tiene dominio $\mathscr{D}_{\mathbf{f}}$ y regla de correspondencia $|\mathbf{f}|(t) = |\mathbf{f}(t)|$.

4. Si **f** tiene la propiedad de que $|\mathbf{f}(t) - \mathbf{f}(s)| \leqslant |t - s|$ para toda t y s en $\mathscr{D}_{\mathbf{f}}$ demuéstrese que **f** es una función continua.

5. CURVAS

El término "curva" tiene significados distintos en distintas áreas de la matemática. Aquí le asignaremos un significado apropiado a nuestro estudio de las funciones vectoriales. Una posibilidad es la de definir una curva como el rango de una función vectorial continua que tiene como dominio un intervalo. Nosotros llamaremos a esto una *curva punteada.* Esta definición es adecuada para la geometría analítica. Según los ejemplos y problemas de la sección 2 vemos que una recta, una circunferencia, una parábola, una elipse y una rama de una hipérbola son ejemplos de curvas punteadas.

Si g es una función real continua con un intervalo $\mathscr{J}$ como dominio, entonces, si hacemos $\mathbf{f} = (I, g)$ vemos que la gráfica de g, $\{(t, g(t)) \mid t \in \mathscr{J}\}$ es el rango de **f** y, por tanto, puede considerarse como una curva punteada en R^2. Sin embargo, cuando se discuten las tangentes a la gráfica se hace uso de la descripción analítica de ésta. Así pues, en este contexto la gráfica es más que solamente un conjunto de puntos. Es un conjunto de puntos trazado de la forma descrita por la función $\mathbf{f} = (I, g)$; es decir, $\mathbf{f}(t)$ va

trazando el conjunto de puntos de izquierda a derecha a medida que t aumenta sobre el intervalo $\mathscr{J}$.

Consideremos ahora el problema de describir el movimiento de una partícula que se mueve en el espacio durante un intervalo de tiempo $[a, b]$. Con cada punto t de $[a, b]$ asociamos el punto $\mathbf{f}(t)$ que es la posición de la partícula en ese instante en relación con cierto sistema de coordenadas rectangulares. De esta forma el movimiento de la partícula queda descrito por la función vectorial $\mathbf{f}$ de dominio $[a, b]$ y rango en R^3. Además, la función $\mathbf{f}$ será continua, pues en mecánica clásica suponemos que una partícula no puede cambiar instantáneamente de posición; es decir, si la partícula está en el punto $\mathbf{P}_0 = \mathbf{f}(t_0)$ en el instante t_0 y $\mathscr{S}(\mathbf{P}_0; \varepsilon)$ es una vecindad de $\mathbf{P}_0$ entonces existe un intervalo de tiempo $\langle t_0 - \delta, t_0 + \delta \rangle$ durante el cual la partícula permanece en la vecindad $\mathscr{S}(\mathbf{P}_0; \varepsilon)$.

En problemas tales como éste, la curva de puntos que es el rango de $\mathbf{f}$ no nos da una descripción adecuada del movimiento de la partícula. Claramente la misma curva de puntos puede haber sido trazada de modos muy diferentes; en diferentes direcciones y con diferentes velocidades. Para describir la forma en que se ha trazado la trayectoria de la partícula tenemos que conocer cuál es la función $\mathbf{f}$, no sólo su rango.

Definimos por ello una *curva-trayectoria* como una función vectorial continua con un intervalo como dominio. En este capítulo trataremos casi exclusivamente de curvas-trayectoria y, por ello, emplearemos simplemente el término "curva" para indicar "curva-trayectoria". Por tanto, una curva es una función $\mathbf{f}$. Sin embargo, como el téımino "curva" debe tener una connotación geométrica, pensamos en una curva como la curva punteada correspondiente (el rango de $\mathbf{f}$) trazada en la forma que $\mathbf{f}$ prescribe. Denotaremos la curva por $\mathscr{C}$ y diremos que $\mathscr{C}$ es la curva descrita por $\mathbf{f}$.

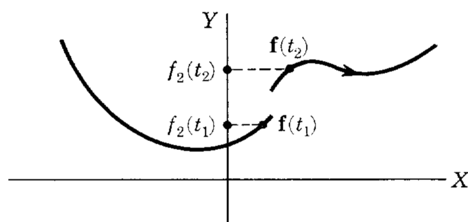


FIGURA 3

Como una curva está descrita por una función continua, no puede haber interrupciones en su trazo. Por ejemplo, el conjunto dibujado en la figura 3 no es una curva de acuerdo con la definición que hemos aceptado. Supongamos que este conjunto fuera una curva descrita por la función continua $\mathbf{f} = (f_1, f_2)$ con el intervalo $\mathscr{J}$ como dominio. Entonces f_1 y f_2 serían continuas sobre $\mathscr{J}$. De acuerdo con el teorema del valor intermedio

para funciones reales,[1] f_1 y f_2 transforman intervalos sobre intervalos. Sin embargo, en la figura 3 vemos que f_2 no transforma el intervalo $[t_1, t_2]$ sobre un intervalo.

5.1 Ejemplo. Proporciónese una descripción geométrica de la curva $\mathscr{C}$ descrita por **f**, donde $f(t) = (\cos t, \operatorname{sen} t)$ y $\mathscr{D}_{\mathbf{f}} = [0, 2\pi]$.

SOLUCIÓN. La curva-punto que es el rango de **f** es la circunferencia de radio uno con centro en el origen: $\{(x, y) \mid x^2+y^2 = 1\}$. A medida que t va de 0 a 2π, el punto $\mathbf{f}(t)$ va recorriendo $\mathscr{C}$, la circunferencia, en dirección contraria a la de las manecillas del reloj desde $\mathbf{f}(0) = (1, 0)$ hasta $\mathbf{f}(2\pi) = (1, 0)$.

La curva $\mathscr{C}$ del ejemplo 5.1 puede también describirse por

$$x = \cos t,\ y = \operatorname{sen} t,\ t \in [0, 2\pi].$$

La variable t se llama parámetro, y estas ecuaciones se llaman ecuaciones paramétricas de $\mathscr{C}$.[2]

5.2 Ejemplo. Trácese la curva dada por las ecuaciones paramétricas:

$$x = \cos t,\ y = \operatorname{sen} t,\ z = \tfrac{1}{2}t,\ t \in [0, 4\pi].$$

SOLUCIÓN. La distancia del eje Z a un punto (x, y, z) cualquiera de la curva es $\sqrt{x^2+y^2} = \sqrt{\cos^2+\operatorname{sen}^2 t} = 1$. Así pues, la curva debe encontrarse sobre el cilindro circular recto con base de radio 1 y el eje Z como eje (figura 4). A medida que t aumenta de 0 a 4π, el punto $(x, y, z) = (\cos t, \operatorname{sen} t, \frac{1}{2}t)$ se mueve de $(1, 0, 0)$ a $(1, 0, 2\pi)$ girando en dirección contraria a la de las manecillas del reloj (cuando se ve desde arriba) y moviéndose hacia arriba sobre la superficie del cilindro. Esta curva es un arco de hélice cilíndrica.

Nota. Si usamos los vectores unitarios $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$ y $\mathbf{k} = (0, 0, 1)$, la regla de correspondencia para la función **f** que describe la curva en el ejemplo 5.2 puede escribirse

$$\mathbf{f}(t) = \cos t\,\mathbf{i} + \operatorname{sen} t\,\mathbf{j} + \tfrac{1}{2}t\,\mathbf{k}.$$

[1] Volumen 1, pág. 430.

[2] El lector puede ver que, según esto, ecuaciones paramétricas de una curva-trayectoria son las funciones componentes de la curva. En cuanto a parámetro, parece que el autor llama aquí así a la letra empleada para representar un elemento no determinado del dominio de la curva, pero esto no encaja bien con el uso de esta palabra en expresiones tales como, por ejemplo, "cambio de parámetro". [N. del T.]

$$\mathbf{f}(0) = (1, 0, 0)$$
$$\mathbf{f}\left(\frac{\pi}{2}\right) = \left(0, 1, \frac{\pi}{4}\right)$$
$$\mathbf{f}(\pi) = \left(-1, 0, \frac{\pi}{2}\right)$$
$$\mathbf{f}\left(\frac{3\pi}{2}\right) = \left(0, -1, \frac{3\pi}{4}\right)$$
$$\mathbf{f}(2\pi) = (1, 0, \pi)$$
$$\mathbf{f}\left(\frac{5\pi}{2}\right) = \left(0, 1, \frac{5\pi}{4}\right)$$
$$\mathbf{f}(3\pi) = \left(-1, 0, \frac{3\pi}{2}\right)$$
$$\mathbf{f}\left(\frac{7\pi}{2}\right) = \left(0, -1, \frac{7\pi}{4}\right)$$
$$\mathbf{f}(4\pi) = (1, 0, 2\pi)$$

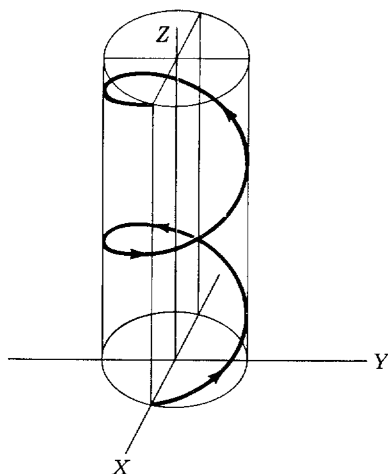


FIGURA 4

5.3 Ejemplo. Provéase una función que tenga como rango la curva punteada trazada por un punto **P** de una circunferencia cuando la circunferencia rueda (sin deslizamiento) sobre una recta. Esta curva punteada se llama cicloide.

SOLUCIÓN. (Figura 5.) Supongamos que la recta sobre la que la circunferencia rueda es el eje X y sea **P** el punto de la circunferencia que se

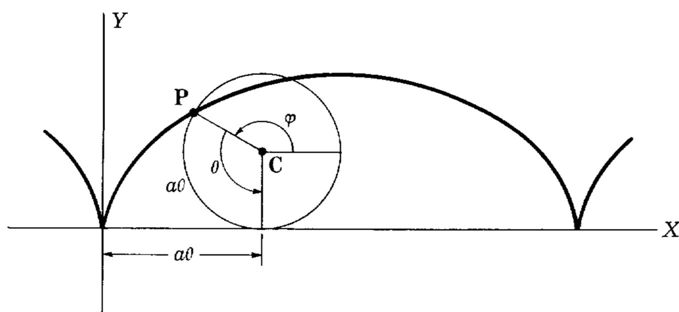


FIGURA 5

encuentra sobre el eje X cuando el centro de la circunferencia está sobre el eje Y. Sea θ la medida en radianes del ángulo que forma el vector $\mathbf{P}-\mathbf{C}$ con la dirección negativa del eje Y y sea φ el ángulo que forma la dirección positiva del eje X con $\mathbf{P}-\mathbf{C}$. Como $\mathbf{C} = (a\theta, a)$ y $\varphi+\theta = \dfrac{3\pi}{2}$, tenemos

$$\begin{aligned}(x, y) &= \mathbf{P} = \mathbf{C}+(a \cos \varphi, a \operatorname{sen} \varphi)\\ &= (a\theta, a)+(-a \operatorname{sen} \theta, -a \cos \theta)\\ &= a(\theta - \operatorname{sen} \theta, 1 - \cos \theta).\end{aligned}$$

Así pues, la cicloide es el rango de la función $\mathbf{f}$ donde

$$\mathbf{f}(\theta) = a(\theta - \operatorname{sen} \theta, 1 - \cos \theta).$$

Problemas

1. Proporciónense descripciones geométricas y dibujos de las curvas descritas por las siguientes funciones:

a) $\mathbf{f}(t) = e^{-t}(\cos 2\pi t, \operatorname{sen} 2\pi t)$
b) $\mathbf{f}(t) = e^{-t}(\operatorname{sen} 2\pi t, \cos 2\pi t)$
c) $\mathbf{f}(t) = (1 - \operatorname{sen} t, -2 + \operatorname{sen} t, 2 \operatorname{sen} t)$.

2. Dibújese la curva descrita por

$$\mathbf{f}(t) = \cos t\mathbf{i}+2 \cos t\mathbf{j} + \operatorname{sen} t\mathbf{k}, \quad t\in[0, 2\pi].$$

3. Dibújese el arco de la hélice cónica descrita por

$$\mathbf{f}(\theta) = \left(\theta \cos \theta, \theta \operatorname{sen} \theta, \frac{\theta}{2\pi}\right), \quad \theta\in[0, 2\pi].$$

4. Proporciónese una función que tenga como rango la curva punteada trazada por un punto $\mathbf{P}$ sobre una circunferencia de radio 1 cuando la circunferencia rueda sobre el lado interior de un círculo de radio 4 y dibújese. A esta curva punteada se le llama hipocicloide.

5. Un punto $\mathbf{P}$ en el primer cuadrante de R^2 se mueve de tal forma que su distancia al origen es igual a la pendiente t de la recta que va del origen a $\mathbf{P}$. Proporciónese una representación paramétrica de la curva trazada por $\mathbf{P}$ usando t como parámetro y dibújese.

6. Sea $\{(x(t), y(t)) \mid t\in\mathrm{R}\}$ la trayectoria de una partícula y denotemos por $\dot{x}$ y $\dot{y}$ las derivadas de las funciones x y y. Determínese la trayectoria si

a) $\ddot{x} = 0$, $\ddot{y} = -g$, $\dot{x}(0) = 1$, $\dot{y}(0) = 2$, $x(0) = 0$, $y(0) = 0$
b) $\dot{x} = x$, $\dot{y} = 2y$, $x(0) = 1$, $y(0) = 2$.

Sugerencia. $\dot{x} = x \Rightarrow D_t(e^{-t}x(t)) = 0$

c) $\dot{x} = x$, $\dot{y} = 2y-x$, $x(0) = 1$, $y(0) = 2$.

6. LA DERIVADA

En el cálculo de funciones reales de una variable real la derivada de una función f se define como la función f' cuya regla de correspondencia es

$$f'(t) = \lim_{h \to 0} \frac{f(t+h)-f(t)}{h}$$

y cuyo dominio es el conjunto de todos los números reales t para los que el anterior límite está definido. Si $\mathbf{f}$ es una función vectorial de una variable real, definimos la derivada de $\mathbf{f}$ en la misma forma.

Nota. Siempre que consideramos derivadas de funciones de una variable suponemos que todas las funciones que aparecen en la discusión están definidas en un intervalo que consta de más de un punto o por la unión de intervalos de tal tipo.

6.1 Definición. *La derivada de una función vectorial* $\mathbf{f}$ *es la función vectorial* $\mathbf{f}'$ *cuya regla de correspondencia es*

$$\mathbf{f}'(t) = \lim_{h \to 0} \frac{\mathbf{f}(t+h)-\mathbf{f}(t)}{h}$$

y cuyo dominio es el conjunto de todos los números reales t *para los que el anterior límite existe.*

Si t es un número en el dominio de $\mathbf{f}'$, entonces se dice que $\mathbf{f}$ es *diferenciable* en t.

Aplicando el teorema 3.3 (pág. 103), obtenemos la siguiente regla para calcular la derivada de una función vectorial: la derivada de la función $\mathbf{f}$ es la función vectorial cuyos componentes son las derivadas de las componentes de $\mathbf{f}$.

6.2 Teorema. *Si* $\mathbf{f} = (f_1, \ldots, f_n)$, *entonces*

$$\mathbf{f}' = (f_1', \ldots, f_n'),$$

donde el dominio de $\mathbf{f}'$ *es la intersección de los dominios de las derivadas* $f_1', \ldots, f_n'$.

PRUEBA. De acuerdo con el teorema 3.3 sabemos que

$$\lim_{h \to 0} \frac{\mathbf{f}(t+h)-\mathbf{f}(t)}{h} = \lim_{h \to 0} \left(\frac{f_1(t+h)-f_1(t)}{h}, \ldots, \frac{f_n(t+h)-f_n(t)}{h} \right)$$

existe si y sólo si cada uno de los límites $\lim_{h \to 0} \frac{f_i(t+h)-f_i(t)}{h}$ $(i = 1, \ldots, n)$

existe. Esto prueba que el dominio de $\mathbf{f}'$ es la intersección de los dominios de $f_1', \ldots, f_n'$. Si t está en el dominio de $\mathbf{f}'$, entonces, usando de nuevo el teorema 3.3, concluimos que

$$\mathbf{f}'(t) = (f_1'(t), \ldots, f_n'(t));$$

es decir, que

$$\mathbf{f}' = (f_1', \ldots, f_n').$$

6.3 Ejemplo. Encuéntrese $\mathbf{f}'$ cuando

a) $\mathbf{f} = (\cos, \operatorname{sen})$
b) $\mathbf{f}(t) = (t, 2-t^3, 4 \ln(1-t)), \ t < 1$
c) $\mathbf{f}(t) = e^{-t}(1, \cos \omega t, \operatorname{sen} \omega t)$.

SOLUCIÓN.

a) $\mathbf{f}' = (-\operatorname{sen}, \cos)$

b) $\mathbf{f}'(t) = \left(1, -3t^2, \dfrac{-4}{1-t}\right), \ t < 1$

c) $\mathbf{f}(t) = (e^{-t}, e^{-t}\cos \omega t, e^{-t} \operatorname{sen} \omega t)$ y

$$\begin{aligned}\mathbf{f}'(t) &= (-e^{-t}, e^{-t}(-\cos \omega t - \omega \operatorname{sen} \omega t), e^{-t}(-\operatorname{sen} \omega t + \omega \cos \omega t)) \\ &= -e^{-t}(1, \cos \omega t, \operatorname{sen} \omega t) + e^{-t}(0, -\omega \operatorname{sen} \omega t, \omega \cos \omega t).\end{aligned}$$

Damos ahora una interpretación geométrica de la derivada de una función vectorial. Sea $\mathscr{C}$ la curva descrita por la función $\mathbf{f}$ cuyo dominio es $\mathscr{J}$. Si t y $t+h$ están en $\mathscr{J}$ $(h \neq 0)$, entonces $\dfrac{1}{h}(f(t+h)-f(t))$ es un vector paralelo a la cuerda que une $\mathbf{f}(t)$ con $\mathbf{f}(t+h)$ (figura 6). Si $\mathbf{f}$ es diferenciable en t y $\mathbf{f}'(t) \neq \mathbf{0}$, entonces la dirección del vector

$$\frac{1}{h}[\mathbf{f}(t+h)-\mathbf{f}(t)]$$

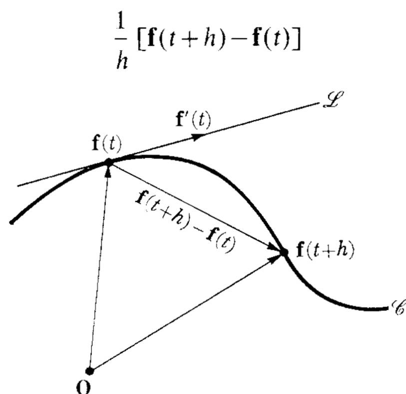


FIGURA 6

se aproxima a la dirección del $\mathbf{f}'(t)$ cuando t tiende a cero, puesto que

$$\mathbf{f}'(t) = \lim_{h \to 0} \frac{\mathbf{f}(t+h) - \mathbf{f}(t)}{h}.$$

Es, por tanto, natural dar la siguiente definición.

6.4 Definición. *Si $\mathscr{C}$ es una curva descrita por* $\mathbf{f}$ *y si* $\mathbf{f}'(t)$ *existe y es distinta de cero, entonces* $\mathbf{f}'(t)$ *se llama* **vector tangente** *a la curva $\mathscr{C}$ en el punto* $\mathbf{f}(t)$, *y la recta*

$$\mathscr{L} = \{\mathbf{f}(t) + r\mathbf{f}'(t) \mid r \in \mathrm{R}\}$$

se llama ***recta tangente*** *a la curva $\mathscr{C}$ en* $\mathbf{f}(t)$.

El vector tangente $\mathbf{f}'(t)$ apunta en la dirección en que la curva va siendo trazada en $\mathbf{f}(t)$ cuando t aumenta.

El siguiente ejemplo muestra que la definición de recta tangente a una curva es una extensión del concepto de recta tangente a la gráfica de una función real de variable real.

6.5 Ejemplo. Si $\mathscr{C}$ es la gráfica de la función real g, demuéstrese que $g'(x)$ es la pendiente de la recta tangente (definida en 6.4) en el punto $(x, g(x))$ de $\mathscr{C}$.

SOLUCIÓN. $\mathscr{C}$ es la curva descrita por la función $\mathbf{f} = (I, g)$. Luego $\mathbf{f}' = (1, g')$ y la recta tangente en el punto $(x, g(x))$ de $\mathscr{C}$ es

$$\mathscr{L} = \{(x, g(x)) + r(1, g'(x)) \mid r \in \mathrm{R}\}.$$

La pendiente de $\mathscr{L}$ es $g'(x)$.

Introducimos ahora otra notación para la derivada. Si la curva está descrita por la transformación $\mathbf{f}$ del intervalo $\mathscr{I}$, entonces $\mathscr{C} = \{\mathbf{x} \mid \mathbf{x} = \mathbf{f}(t) t, \in \mathscr{I}\}$ y decimos que $\mathscr{C}$ está descrita por la ecuación paramétrica

$$\mathbf{x} = \mathbf{f}(t).$$

Sea

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}'(t).$$

Si $\mathscr{C}$ es una curva en el espacio tridimensional, entonces tiene una ecuación

$$\mathbf{x} = (x, y, z) = \mathbf{f}(t)$$

y

$$\frac{d\mathbf{x}}{dt} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) = \mathbf{f}'(t).$$

6.6 Ejemplo. Encuéntrese la recta tangente a la hélice cilíndrica (figura 4, pág. 113) descrita por las ecuaciones paramétricas

$$x = \cos t$$
$$y = \operatorname{sen} t$$
$$z = \tfrac{1}{2} t, \qquad t \in \langle -\infty, \infty \rangle$$

en el punto $\left(0, 1, \dfrac{\pi}{4}\right)$.

SOLUCIÓN. Nótese que el punto $\left(0, 1, \dfrac{\pi}{4}\right)$ corresponde a $t = \dfrac{\pi}{2}$. La curva $\mathscr{C}$ está descrita por la ecuación

$$(x, y, z) = (\cos, t, \operatorname{sen} t, \tfrac{1}{2} t) = \mathbf{f}(t).$$

Entonces

$$\mathbf{f}'(t) = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) = (-\operatorname{sen} t, \cos t, \tfrac{1}{2})$$

y

$$\mathbf{f}'\left(\frac{\pi}{2}\right) = (-1, 0, \tfrac{1}{2}).$$

Por tanto, la recta tangente a $\mathscr{C}$ en $\left(0, 1, \dfrac{\pi}{4}\right)$ es

$$\mathscr{L} = \left\{\left(0, 1, \frac{\pi}{4}\right) + r(-1, 0, \tfrac{1}{2}) \mid r \in \mathrm{R}\right\}$$

y son ecuaciones paramétricas de $\mathscr{L}$

$$x = -r$$
$$y = 1$$
$$z = \frac{\pi}{4} + \frac{1}{2} r, \quad r \in \langle -\infty, \infty \rangle.$$

Los siguientes ejemplos ilustran la manera en que la derivada puede usarse como una ayuda para el dibujo de la curva.

6.7 Ejemplo. Dibújese la curva $\mathscr{C}$ descrita por

$$\mathbf{f} = (l^3 - 4l, l^2 - 4).$$

SOLUCIÓN. Como $f_1 = t^3-4t$ es una función impar y $f_2 = t^2-4$ es una función par, la curva es simétrica respecto al eje Y; si $\mathbf{f}(t_0) = (x_0, y_0)$ entonces $\mathbf{f}(-t_0) = (-x_0, y_0)$. Podemos pues restringir nuestra atención a valores no negativos de t. Como $\mathbf{f}' = (3t^2-4, 2t)$, la curva tiene un vector tangente $\mathbf{f}'(t) = (3t^2-4, 2t)$ en cada punto $\mathbf{f}(t)$. Al dibujar $\mathscr{C}$ los puntos donde $\mathbf{f}'(t)$ es horizontal (con segunda componente cero) o vertical (con primera componente cero) son de interés particular. En $\mathbf{f}(0) = (0, -4)$ tiene un vector tangente horizontal $\mathbf{f}'(0) = (-4, 0)$ y en $\mathbf{f}\left(\dfrac{2}{\sqrt{3}}\right) = \left(-\dfrac{16}{9}\sqrt{3}, -\dfrac{8}{3}\right)$ la curva tiene un vector tangente vertical $\mathbf{f}'\left(\dfrac{2}{\sqrt{3}}\right) = \left(0, \dfrac{4}{3}\sqrt{3}\right)$. Considerando, además, la expresión general del vector tangente

$$\mathbf{f}'(t) = (3t^2-4, 2t),$$

tenemos:

Si $t \in \langle 0, \frac{2}{3}\sqrt{3}\rangle$, entonces $\mathbf{f}'(t)$ apunta hacia la izquierda y hacia arriba puesto que $3t^2-4$ es negativa y $2t$ es positiva.

Si $t \in \langle \frac{2}{3}\sqrt{3}, \infty\rangle$, entonces $\mathbf{f}'(t)$ apunta a la derecha y hacia arriba puesto que $3t^2-4$ y $2t$ son positivas.

Marcando ahora algunos puntos (entre los que deben incluirse todas las intersecciones con los ejes coordenados) podemos dibujar $\mathscr{C}$ (figura 7). El punto (0, 0) se llama punto doble de $\mathscr{C}$: $\mathbf{f}(-2) = \mathbf{f}(2) = (0, 0)$. Nótese que $\mathscr{C}$ tiene dos vectores tangentes en este punto: $\mathbf{f}'(-2) = (8, -4)$ y $\mathbf{f}'(2) = (8.4)$.

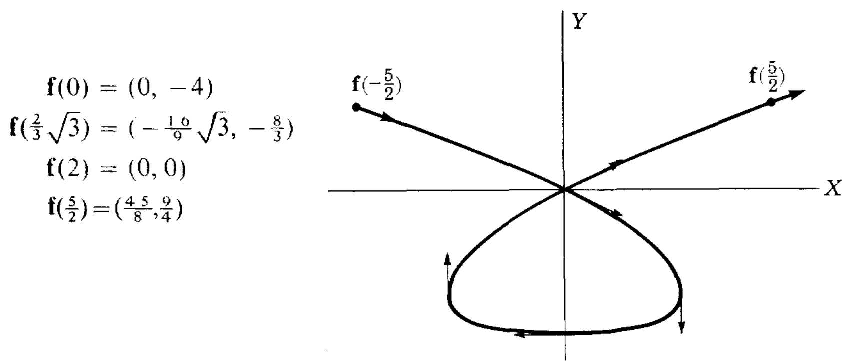


FIGURA 7

Obsérvese que en la definición 6.4 no se define ningún vector tangente en el punto $\mathbf{f}(t)$ si $\mathbf{f}'(t) = 0$. En tal punto puede suceder que la curva tenga un cambio de dirección abrupto. Ilustramos esto en el siguiente ejemplo.

6.8 Ejemplo. Dibújese la curva $\mathscr{C}$ descrita por

$$\mathbf{f}(t) = \left(\frac{t^2}{1+t^2}, \frac{t^3}{1+t^2}\right), \quad t \in \langle -\infty, \infty \rangle .$$

SOLUCIÓN. Como f_1 es una función par y f_2 es una función impar, $\mathscr{C}$ es simétrica con respecto al eje X; si $\mathbf{f}(t_0) = (x_0, y_0)$ entonces $\mathbf{f}(-t_0) = (x_0, -y_0)$. Considerando el vector tangente

$$\mathbf{f}'(t) = \left(\frac{2t}{(1+t^2)^2}, \frac{t^4+3t^2}{(1+t^2)^2}\right),$$

vemos que $\mathscr{C}$ no tiene tangentes horizontales ni verticales. Sin embargo $\mathbf{f}'(0) = \mathbf{0}$. Investigamos ahora el comportamiento de $\mathscr{C}$ en el punto $\mathbf{f}(0) = (0, 0)$. Escribiendo

$$\mathbf{f}'(t) = \frac{t}{(1+t^2)^2}(2, t^3+3t),$$

vemos que, para $t < 0$, $\mathbf{f}'(t)$ tiene la misma dirección que $-(2, t^3+3t)$ y, para $t > 0$, $\mathbf{f}'(t)$ tiene la misma dirección que $(2, t^3+3t)$. Como

$$\lim_{t \to 0^-} -(2, t^3+3t) = (-2, 0) \quad \text{y} \quad \lim_{t \to 0^+} (2, t^3+3t) = (2, 0),$$

la curva tiene un abrupto cambio de dirección en $\mathbf{f}(0)$ (figura 8). A tal punto se le llama cúspide o punto cuspidal. La recta $x = 1$ es una asíntota vertical de $\mathscr{C}$:

$$\lim_{t \to \infty} \frac{t^2}{1+t^2} = 1 \quad \text{y} \quad \lim_{t \to \infty} \frac{t^3}{1+t^2} = \infty .$$

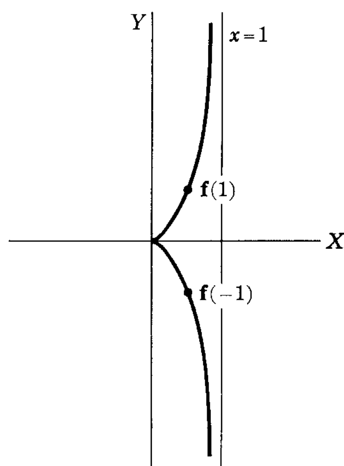


FIGURA 8

Si una función $\mathbf{f}$ describe el movimiento de una partícula durante un intervalo de tiempo $\mathscr{J}$ [es decir, para cualquier $t \in \mathscr{J}$, $\mathbf{f}(t)$ es la posición de la partícula en el tiempo t], entonces $\mathbf{f}'(t)$ es la *velocidad* y $|\mathbf{f}'(t)|$ es la *rapidez* o *velocidad modular* de la partícula en el instante t.

Así pues, $\mathbf{f}'(t) = \mathbf{0}$ significa que la partícula tiene velocidad cero en el instante t. Como hemos visto, puede que haya un cambio abrupto de dirección en la trayectoria en el punto $\mathbf{f}(t)$. Sin embargo, no es este necesariamente el caso. Por ejemplo, supongamos $\mathbf{f} = (I^3, I^3)$. Entonces $\mathbf{f}' = (3I^2, 3I^2)$ y $\mathbf{f}'(0) = \mathbf{0}$. La curva descrita por $\mathbf{f}$ es la recta con ecuación $y = x$ trazada de izquierda a derecha cuando t aumenta. El hecho de que $\mathbf{f}'(0) = \mathbf{0}$ significa que la partícula se para en el origen.

Problemas

1. Determínese $\mathbf{f}'$ cuando

a) $\mathbf{f} = (I^{1/3}, 3I^2, \text{sen})$

b) $\mathbf{f} = (\exp, \text{senh}, \cosh)$

c) $\mathbf{f}(t) = \left(\ln(t^2+1), \sqrt{t^2+1}, \dfrac{2t}{t^2+1}\right)$

d) $\mathbf{f}(t) = (\cos^2 3t, \cos t \,\text{sen}\, t, \tan t)$

e) $\mathbf{f}(t) = (t, t, [t])$

f) $\begin{cases} \mathbf{f}(t) = \left(e^{2t}, t^2 \,\text{sen}\, \dfrac{1}{t}\right), & t \neq 0 \\ \mathbf{f}(0) = (1, 0). \end{cases}$

2. Pruébese que si $\mathbf{f}$ es diferenciable en el punto t, entonces $\mathbf{f}$ es continua en t.

3. Encuéntrese un vector tangente y la recta tangente a

a) la elipse con ecuaciones paramétricas

$$x = 4\cos\theta$$
$$y = 3\,\text{sen}\,\theta,\ \theta \in [0, 2\pi]$$

en los puntos $(0, 3)$, $(2\sqrt{2}, \tfrac{3}{2}\sqrt{2})$, $(4, 0)$;

b) la recta de representación paramétrica

$$\mathbf{f}(t) = (5, -3, 8) + t(8, -17, 32)$$

en los puntos $(5, -3, 8)$, $(13, -20, 40)$, $(29, -54, 104)$;

c) la hélice cónica de representación paramétrica

$$\mathbf{f}(\theta) = \left(\theta\cos\theta, \theta\,\text{sen}\,\theta, \frac{\theta}{2\pi}\right)$$

en los puntos $(0, 0, 0)$, $\left(0, \dfrac{\pi}{2}, \dfrac{1}{4}\right)$.

4. Trácese la curva $\mathscr{C}$ descrita por la función **f** en cada uno de los casos que siguen. Encuéntrense todos los puntos en que $\mathscr{C}$ tiene un vector tangente horizontal o vertical.

a) $\mathbf{f} = (I^3, I^2+2I)$

b) $\mathbf{f} = (I^4-4I, I^3)$

c) $\mathbf{f}(t) = (\cos t, \operatorname{sen} 3t), t \in [0, 2\pi]$

d) $\mathbf{f}(t) = (\cos 2t, \cos 2t \tan t), t \in \left[-\frac{\pi}{3}, \frac{\pi}{3}\right].$

5. Trácese la curva $\mathscr{C}$ descrita por la función **f** en cada uno de los casos que siguen. Encuéntrense todos los puntos en que $\mathscr{C}$ tiene un vector tangente paralelo a uno de los planos coordenados.

a) $\mathbf{f}(t) = (\operatorname{sen} t, \cos t, \operatorname{sen} 3t)$
b) $\mathbf{f}(t) = (\operatorname{sen} 2t, \cos t, \operatorname{sen} 3t).$

6. Demuéstrese que la cicloide descrita por $\mathbf{f} = a(I - \operatorname{sen}, 1 - \cos)$ tiene un punto cuspidal en los puntos $\mathbf{f}(0) = (0, 0)$ y $\mathbf{f}(2\pi) = (2\pi a, 0)$. (Véase la figura 5, pág. 113.)

7. Trácese la curva $\mathscr{C}$ descrita por la función **f** en cada uno de los siguientes casos. Determínense todos los puntos de $\mathscr{C}$ en que $\mathbf{f}'(t) = \mathbf{0}$ y discútase el comportamiento de $\mathscr{C}$ en estos puntos.

a) $\mathbf{f}(t) = (t^2, t^3), t \in [-1, 1]$
b) $\mathbf{f}(t) = (t^3, t^5), t \in [-1, 1]$
c) $\mathbf{f}(t) = (t^3, t^3, |t^3|), t \in [-1, 1]$
d) $\mathbf{f}(t) = (t^4 - 2t^2, t^3), t \in [-1, 1].$

8. Si **f** está definida sobre $[a, b]$ y $\mathbf{f}'(t) = \mathbf{0}$ para todo $t \in [a, b]$, demuéstrese que **f** es una constante sobre $[a, b]$.

9. Supongamos que tenemos un espejo parabólico formado por la rotación alrededor de su eje de la parábola descrita por

$$\mathbf{f}(\theta) = \frac{d}{1 - \cos\theta}(\cos\theta, \operatorname{sen}\theta).$$

Demuéstrese que un rayo de luz que emana del foco de la parábola se refleja paralelamente al eje.

7. ALGUNOS TEOREMAS SOBRE LA DERIVADA

Se dice que una función es diferenciable en un punto si la derivada de la función existe en tal punto. Definimos ahora lo que significa decir que una función es diferenciable en un intervalo. La función **f** es *diferenciable sobre el intervalo abierto* $\langle a, b\rangle$ si **f** es diferenciable en cada punto de $\langle a, b\rangle$, y la función **f** es *diferenciable sobre el intervalo cerrado* $[a, b]$ si **f** es diferenciable sobre el intervalo abierto $\langle a, b\rangle$ y si existen las siguientes derivadas laterales en los puntos extremos:

$$\mathbf{f}'^{+}(a) = \lim_{h\to 0^{+}} \frac{\mathbf{f}(a+h)-\mathbf{f}(a)}{h}$$

y

$$\mathbf{f}'^{-}(b) = \lim_{h\to 0^{-}} \frac{\mathbf{f}(b+h)-\mathbf{f}(b)}{h}.$$

El siguiente teorema es una simple consecuencia de las definiciones de continuidad y diferenciabilidad sobre un intervalo.

7.1 Teorema. *Si la función* **f** *es diferenciable sobre un intervalo* $\mathscr{J}$, *entonces* **f** *es continua sobre* $\mathscr{J}$.

En el cálculo de funciones vectoriales hay reglas para el cálculo de derivadas que son análogas a las existentes para funciones reales; por ejemplo, la derivada de una suma es la suma de las derivadas. Antes de dar estas reglas introduciremos una notación que nos permita formularlas de un modo conveniente. Hagamos $\mathbf{f}' = D\mathbf{f}$; D es una función (operador) cuyo valor en **f** es **f**′. Las funciones cuyo dominio y rango son conjuntos de funciones se llaman usualmente operadores. De ahora en adelante diremos que la función **f**′ se obtiene cuando aplicamos el operador D a **f**.

7.2 Teorema. *Si* **f**, **g** *y* φ *son diferenciables sobre un intervalo* $\mathscr{J}$, *entonces* $\mathbf{f}+\mathbf{g}$, $\mathbf{f}-\mathbf{g}$, $\mathbf{f}\cdot\mathbf{g}$, $\mathbf{f}\times\mathbf{g}$ *y* $\varphi\mathbf{f}$ *son diferenciables sobre* $\mathscr{J}$ *y, sobre* $\mathscr{J}$,

$$\begin{aligned}
D(\mathbf{f}+\mathbf{g}) &= D\mathbf{f}+D\mathbf{g}\\
D(\mathbf{f}-\mathbf{g}) &= D\mathbf{f}-D\mathbf{g}\\
D(\mathbf{f}\cdot\mathbf{g}) &= \mathbf{f}\cdot D\mathbf{g}+D\mathbf{f}\cdot\mathbf{g}\\
D(\mathbf{f}\times\mathbf{g}) &= \mathbf{f}\times D\mathbf{g}+D\mathbf{f}\times\mathbf{g}\\
D(\varphi\mathbf{f}) &= \varphi(D\mathbf{f})+(D\varphi)\mathbf{f}.
\end{aligned}$$

PRUEBA. Probaremos solamente la parte del teorema sobre el producto vectorial. Las pruebas de las restantes fórmulas son análogas. Si $\mathbf{f} = (f_1, f_2, f_3)$ y $\mathbf{g} = (g_1, g_2, g_3)$, entonces

$$\mathbf{f}\times\mathbf{g} = (f_2g_3-f_3g_2, f_3g_1-f_1g_3, f_1g_2-f_2g_1).$$

Por el teorema 6.2 (pág. 00), sobre el intervalo $\mathscr{J}$,

$$\begin{aligned}
D(\mathbf{f}\times\mathbf{g}) &= (D(f_2g_3-f_3g_2),\ D(f_3g_1-f_1g_3),\ D(f_1g_2-f_2g_1)) \\
&= (f_2\,Dg_3+g_3\,Df_2-f_3\,Dg_2-g_2\,Df_3,\ f_3\,Dg_1+g_1\,Df_3-f_1\,Dg_3 \\
&\qquad -g_3\,Df_1,\ f_1\,Dg_2+g_2\,Df_1-f_2\,Dg_1-g_1\,Df_2) \\
&= (f_2\,Dg_3-f_3\,Dg_2,\ f_3\,Dg_1-f_1\,Dg_3,\ f_1\,Dg_2-f_2\,Dg_1) \\
&\qquad +(g_3\,Df_2-g_2\,Df_3,\ g_1\,Df_3-g_3\,Df_1,\ g_2\,Df_1-g_1\,Df_2) \\
&= (f_1, f_2, f_3)\times(Dg_1,\ Dg_2,\ Dg_3)+(Df_1,\ Df_2,\ Df_3)\times(g_1, g_2, g_3) \\
&= \mathbf{f}\times D\mathbf{g}+D\mathbf{f}\times\mathbf{g}.
\end{aligned}$$

Nota. Como el producto vectorial no es conmutativo, se debe tener cuidado en escribir los factores en la fórmula para la derivada del producto vectorial en el orden correcto. Esta fórmula es para funciones vectoriales con rango en R^3, solamente; todas las otras fórmulas se verifican para funciones vectoriales con rango en un espacio euclidiano de una dimensión finita cualquiera.

7.3 Ejemplo. Demuéstrese que si $|\mathbf{f}|$ es una constante, entonces $\mathbf{f}(t)$ y $\mathbf{f}'(t)$ son ortogonales para todo $t\in\mathscr{D}_{\mathbf{f}'}$.

SOLUCIÓN. Para todo $t\in\mathscr{D}_{\mathbf{f}}$

$$\mathbf{f}(t)\cdot\mathbf{f}(t) = |\mathbf{f}(t)|^2 = |\mathbf{f}|^2\,(t).$$

Por tanto, si $|\mathbf{f}| = c$

$$\mathbf{f}\cdot\mathbf{f} = |\mathbf{f}|^2 = c^2$$

y

$$D(\mathbf{f}\cdot\mathbf{f}) = \mathbf{f}\cdot D\mathbf{f}+D\mathbf{f}\cdot\mathbf{f} = 2\mathbf{f}\cdot D\mathbf{f} = 0.$$

Así pues, para todo $t\in\mathscr{D}_{\mathbf{f}'}$, $\mathbf{f}(t)\cdot\mathbf{f}'(t) = 0$; es decir, $\mathbf{f}(t)$ y $\mathbf{f}'(t)$ son ortogonales.

Los símbolos "D_t" y "d/dt" están también en uso para denotar la derivación:

$$D_t\mathbf{f}(t) = \frac{d}{dt}\mathbf{f}(t) = \frac{d\mathbf{f}}{dt} = \mathbf{f}'(t);$$

es decir, si $\mathbf{f}(t)$ es la regla de correspondencia para $\mathbf{f}$, entonces $D_t\mathbf{f}(t) = \dfrac{d}{dt}\mathbf{f}(t)$ denota la regla de correspondencia para $\mathbf{f}'$.[1]

Sea $\mathbf{f}$ una función diferenciable que describe la circunferencia $\mathscr{C}(\mathbf{P}_0; r)$ con centro en $\mathbf{P}_0$ y radio r. Para todo $t\in\mathscr{D}_{\mathbf{f}}$, $|\mathbf{f}(t)-\mathbf{P}_0| = r$ y, por tanto, de acuerdo con el ejemplo 7.3, $\mathbf{f}(t)-\mathbf{P}_0$ es ortogonal a

$$D_t[\mathbf{f}(t)-\mathbf{P}_0] = D_t\mathbf{f}(t)-D_t\mathbf{P}_0 = \mathbf{f}'(t);$$

[1] Es más frecuente decir que $\mathbf{f}(t)$ denota la imagen de t según $\mathbf{f}$, etc., que decir que $\mathbf{f}(t)$ es la regla de correspondencia para $\mathbf{f}$, etc. [N. del T.]

es decir, el radio trazado desde $\mathbf{P}_0$ al punto $\mathbf{f}(t)$ sobre la circunferencia es ortogonal al vector tangente en este punto.

7.4 Ejemplo. Si $\mathbf{f} = (I, \cos, \text{sen})$ y $\varphi = \exp \circ 2I$, determínese $D(\varphi\mathbf{f})$.

Antes de dar la solución de 7.4 repasaremos la definición de composición de funciones reales de una variable real. La función $f \circ g$ (lo que se lee: "f composición g" o "f círculo g") es la función con regla de correspondencia $[f \circ g](x) = f(g(x))$ y el conjunto $\{x \in \mathscr{D}_g \mid g(x) \in \mathscr{D}_f\}$ como dominio. Por tanto,

$$\varphi(t) = [\exp \circ 2I](t) = \exp(2t) = e^{2t}$$

y $\mathscr{D}_\varphi = \text{R}$. La fórmula para la derivada de $f \circ g$, llamada regla de la cadena, es

$$(f \circ g)' = (f' \circ g)g'.$$

Por tanto, $D(\exp \circ 2I) = (\exp \circ 2I)2$; es decir, $\varphi'(t) = 2e^{2t}$.

SOLUCIÓN DE 7.4. Las funciones $\mathbf{f}$ y φ son diferenciables sobre R y

$$\begin{aligned} D(\varphi\mathbf{f}) &= \varphi(D\mathbf{f}) + (D\varphi)\mathbf{f} \\ &= [\exp \circ 2I](1, -\text{sen}, \cos) + [2\exp \circ 2I](I, \cos, \text{sen}) \\ &= [\exp \circ 2I](1 + 2I, 2\cos - \text{sen}, \cos + 2\,\text{sen}); \end{aligned}$$

es decir, para todo $t \in \text{R}$,

$$D_t[\varphi\mathbf{f}](t) = e^{2t}(1 + 2t, 2\cos t - \text{sen}\, t, \cos t + 2\,\text{sen}\, t).$$

Definimos ahora la composición de una función vectorial $\mathbf{f}$ con una función real φ y pasamos a estudiar algunas de las propiedades de esta composición.

7.5 Definición. *Si φ es una función real de variable real y $\mathbf{f}$ es una función vectorial de variable real, $\mathbf{f} \circ \varphi$ es la función vectorial de variable real con regla de correspondencia*

$$[\mathbf{f} \circ \varphi](t) = \mathbf{f}(\varphi(t))$$

y dominio $\mathscr{D}_{\mathbf{f}\circ\varphi} = \{t \in \mathscr{D}_\varphi : \varphi(t) \in \mathscr{D}_{\mathbf{f}}\}$.

Si $t \in \mathscr{D}_{\mathbf{f}\circ\varphi}$ y $\mathbf{f} = (f_1, \ldots, f_n)$, entonces

$$\begin{aligned} [\mathbf{f} \circ \varphi](t) &= \mathbf{f}(\varphi(t)) = (f_1(\varphi(t)), \ldots, f_n(\varphi(t))) \\ &= ([f_1 \circ \varphi](t), \ldots, [f_n \circ \varphi](t)). \end{aligned}$$

Por tanto

$$\mathbf{f} \circ \varphi = (f_1 \circ \varphi, \ldots, f_n \circ \varphi).$$

7.6 Teorema. *Si φ es continua en t_0 y* **f** *es continua en $\varphi(t_0)$, entonces* **f** $\circ$ φ *es continua en t_0.*

PRUEBA. De acuerdo con el teorema 4.2, pág. 109, $\mathbf{f} \circ \varphi$ es continua en t_0 si y sólo si $f_i \circ \varphi$ $(i = 1, \ldots, n)$ es continua en t_0. Como φ es continua en t_0 y f_i es continua en $\varphi(t_0)$, sabemos, de acuerdo con la teoría de funciones reales de variables real que $f_i \circ \varphi$ es continua en t_0.[1] Y esto completa la prueba.

7.7 Teorema. *Si φ es diferenciable sobre un intervalo $\mathscr{J}$ y* **f** *es diferenciable sobre un intervalo que contiene a $\varphi(\mathscr{J}) = \{\varphi(t) \mid t \in \mathscr{J}\}$, entonces* **f** $\circ$ φ *es diferenciable sobre $\mathscr{J}$ y*

$$D(\mathbf{f} \circ \varphi) = [(D\mathbf{f}) \circ \varphi] D\varphi \quad \textit{sobre } \mathscr{J}.$$

PRUEBA. Según el teorema 6.2 (pág. 115), sobre el intervalo $\mathscr{J}$,

$$D(\mathbf{f} \circ \varphi) = (D(f_1 \circ \varphi), \ldots, D(f_n \circ \varphi)).$$

De acuerdo con la regla de la cadena para funciones reales de variable real tenemos: para $i = 1, \ldots, n$

$$D(f_i \circ \varphi) = [(Df_i) \circ \varphi] D\varphi \quad \text{sobre } \mathscr{J}.$$

Así pues

$$\begin{aligned} D(\mathbf{f} \circ \varphi) &= ([(Df_1) \circ \varphi] D\varphi, \ldots, [(Df_n) \circ \varphi] D\varphi) \\ &= ((Df_1) \circ \varphi, \ldots, (Df_n) \circ \varphi) D\varphi \\ &= [(D\mathbf{f}) \circ \varphi] D\varphi. \end{aligned}$$

Y esto completa la prueba.

Podemos escribir la fórmula del teorema 7.7 en la forma

$$D_t \mathbf{f}(\varphi(t)) = \varphi'(t) \mathbf{f}'(\varphi(t)).$$

Un teorema importante en el cálculo de funciones reales de variable real es el teorema del valor medio:

7.8 Si f es continua sobre $[a, b]$ y diferenciable sobre $\langle a, b \rangle$ entonces hay un punto $c \in \langle a, b \rangle$ tal que

$$f(b) - f(a) = (b - a) f'(c).$$

La generalización del teorema del valor medio para funciones vectoriales es la siguiente:

7.9 Teorema. *Si* **f** *es continua sobre $[a, b]$ y es diferenciable sobre $\langle a, b \rangle$ entonces existen $c_i \in \langle a, b \rangle$ tales que*

$$\mathbf{f}(b) - \mathbf{f}(a) = (b - a)(f_1'(c_1), \ldots, f_n'(c_n)).$$

[1] Volumen I, pág. 367.

PRUEBA. La hipótesis sobre **f** implica que cada componente f_i, es una función continua sobre $[a, b]$ y diferenciable sobre $\langle a, b\rangle$. La conclusión del teorema sigue de la aplicación de 7.8 a cada componente f_i de **f**.

Nota. El teorema 5 muestra que bajo las hipótesis del teorema 7.9 no podemos concluir que $\mathbf{f}(b)-\mathbf{f}(a) = (b-a)\,\mathbf{f}'(c)$ para alguna $c\in\langle a, b\rangle$.

Problemas

1. Si

$$\begin{aligned} \mathbf{f}(t) &= (t, t^2, \tfrac{1}{3}t^3),\ t\in[0, \infty\rangle \\ \mathbf{g} &= (\cos, \text{sen}, I) \\ \varphi(t) &= e^{-\alpha t},\ t\in[0, \infty\rangle \end{aligned}$$

determínense:

a) $\mathbf{f}'$ *b*) $\mathbf{g}'$ *c*) $\mathbf{f}''(t)$

d) $D_t^{\,2}\mathbf{g}(t)$ *e*) $\mathbf{f}'(0)+\mathbf{g}'\left(\dfrac{\pi}{2}\right)$ *f*) $D(\mathbf{f}+\mathbf{g})$

g) $(\mathbf{f}\cdot\mathbf{g})'$ *h*) $D(\mathbf{f}\times\mathbf{g})$ *i*) $(\varphi\mathbf{f})'$

j) $\varphi\mathbf{f}'$ *k*) $D(\mathbf{f}\circ\varphi)$ *l*) $(\mathbf{f}\circ\text{sen})'$

m) $\dfrac{d}{dt}\mathbf{g}(t^2)$ *n*) $D(|\mathbf{f}|^2)$ *o*) $D(|\mathbf{f}|)$.

2. Determínense

a) $D_t(a\cos\omega t, a\,\text{sen}\,\omega t)$ *b*) $D_t^{\,2}(a\cos\omega t, a\,\text{sen}\,\omega t)$.

3. ¿Cuál es el dominio y regla de correspondencia para

a) $D(|\mathbf{f}|)$ *b*) $D\left(\dfrac{\mathbf{f}}{|\mathbf{f}|}\right)$?

4. Supongamos que una curva punteada $\mathscr{C}$ está descrita por la función **f** de $[a, b]$ y por la función de $[0, b-a]$ por **g**, donde $\mathbf{g}(u) = \mathbf{f}(b-u)$. ¿Cuál es la relación entre los vectores tangentes determinados por **f** y **g** en cualquier punto de la curva?

5. Consideremos el arco $\mathscr{C}$ de hélice cilíndrica descrito por

$$\mathbf{f}(t) = (\cos t, \text{sen}\, t, t),\ t\in\left[0, \frac{\pi}{2}\right].$$

Demuéstrese que en ningún punto de $\mathscr{C}$ $\mathbf{f}'(t)$ es paralela a la cuerda de $\mathbf{f}(0)$ a $\mathbf{f}\left(\dfrac{\pi}{2}\right)$.

6. Determínese el componente radial [es decir, el componente en la dirección de $\mathbf{f}(t)$] de $\mathbf{f}'(t)$ y de $\mathbf{f}''(t)$, cuando

a) $\mathbf{f}(t) = (r \cos \omega t, r \operatorname{sen} \omega t)$
b) $\mathbf{f}(t) = (r \cos t^2, r \operatorname{sen} t^2)$.

7. La gráfica polar de $r = \theta$ es una espiral de Arquímedes. Son, pues, ecuaciones paramétricas de la espiral de Arquímedes

$$x = \theta \cos \theta$$
$$y = \theta \operatorname{sen} \theta.$$

Determínese un vector tangente a la espiral en el punto $(-\pi, 0)$.

8. Sea g una función real diferenciable sobre $[\alpha, \beta]$ y sea $\mathscr{C}$ la gráfica polar de $r = g(\theta)$. Entonces $\mathscr{C}$ está descrita por la función $\mathbf{f} = g\mathbf{u}$ de $[\alpha, \beta]$ donde $\mathbf{u} = (\cos, \operatorname{sen})$.
Demuéstrese que

$$\mathbf{f}' = g'\mathbf{u} + g\mathbf{u}^{\perp} \quad \text{donde} \quad \mathbf{u}^{\perp} = (-\operatorname{sen}, \cos)$$

e interprétese este resultado geométricamente.

9. Resuélvase el problema 7 usando el problema 8.

10. Determínese un vector tangente en cualquier punto de la cardioide cuya ecuación polar es $r = 1 + \cos \theta$. Dibújese la curva.

11. Supongamos que $\mathbf{f}$ es una función de R a R^2 que es continua sobre $[a, b]$ y diferenciable sobre $\langle a, b \rangle$, donde $a < b$.

a) Demuéstrese que existe un número $c \in \langle a, b \rangle$ tal que

$$[\mathbf{f}(b) - \mathbf{f}(a)]^{\perp} \cdot \mathbf{f}'(c) = 0.$$

Interprétese este resultado geométricamente. Notación: si $\mathbf{a}$ es el vector (a_1, a_2) en R^2, entonces $\mathbf{a}^{\perp}$ denota al vector $(-a_2, a_1)$.

b) Si $f_1(b) - f_1(a) \neq 0$ y, para todo $x \in \langle a, b \rangle$, $\mathbf{f}'(x) \neq \mathbf{0}$, demuéstrese que la fórmula de la parte *a* puede escribirse en la forma

$$\frac{f_2(b) - f_2(a)}{f_1(b) - f_1(a)} = \frac{f_2'(c)}{f_1'(c)}.$$

Esta es la generalización de Cauchy del teorema del valor medio. Se reduce al teorema del valor medio cuando $f_i = I$.

c) Demuéstrese que, con las condiciones de la parte *b*, si $\mathbf{f}(a) = \mathbf{0}$ y $\lim\limits_{x \to a} \dfrac{f_2'(x)}{f_1'(x)}$ existe, entonces

$$\lim_{x \to a} \frac{f_2(x)}{f_1(x)} = \lim_{x \to a} \frac{f_2'(x)}{f_1'(x)}.$$

Se conoce esto como la regla de l'Hospital.

d) Úsese la regla de l'Hospital para evaluar los siguientes límites:

1) $\lim\limits_{x\to 0} \dfrac{e^x - e^{-x}}{x}$ 2) $\lim\limits_{x\to 0} \dfrac{e^x - 1}{\ln(1+x)}$

3) $\lim\limits_{t\to \pi} \dfrac{\tan 3t}{(t-\pi)\cos t}$ 4) $\lim\limits_{x\to 1} \dfrac{\ln(4x^2-3)}{\ln x}$

5) $\lim\limits_{x\to 0} \left(\dfrac{1}{e^x - 1} - \dfrac{1}{x}\right)$.

8. LA DIFERENCIAL

Sea $\mathbf{f}$ una función vectorial definida sobre $[a, b]$ y sean t y $t+h$ puntos distintos en $[a, b]$. El vector $\Delta\mathbf{f}(t; h) = \mathbf{f}(t+h) - \mathbf{f}(t)$ se llama incremento de $\mathbf{f}$ en t correspondiente al incremento h de t; éste es el cambio de $\mathbf{f}$ debido al cambio h en t. Si $\mathbf{f}$ es diferenciable en t, entonces

$$\Delta\mathbf{f}(t; h) = \mathbf{f}(t+h) - \mathbf{f}(t) = h\mathbf{f}'(t) + h\boldsymbol{\varphi}(t; h)$$

donde $\varphi(t; h) = \dfrac{1}{h}[\mathbf{f}(t+h) - \mathbf{f}(t)] - \mathbf{f}'(t)$. Como $\lim\limits_{h\to 0} \varphi(t; h) = \mathbf{0}$, el incremento $\Delta\mathbf{f}(t; h)$ es aproximadamente igual a $h\mathbf{f}'(t)$ para pequeños valores de h. Al término $h\mathbf{f}'(t)$ se le llama diferencial.

8.1 Definición. *El vector $h\mathbf{f}'(t)$ se llama* ***diferencial*** *de $\mathbf{f}$ en t correspondiente al incremento h en t y se denota por $d\mathbf{f}(t; h)$; es decir,*

$$d\mathbf{f}(t; h) = h\mathbf{f}'(t).$$

En términos de la diferencial tenemos

$$\Delta\mathbf{f}(t; h) = d\mathbf{f}(t; h) + h\boldsymbol{\varphi}(t; h)$$

donde $\lim\limits_{h\to 0} \boldsymbol{\varphi}(t; h) = \mathbf{0}$. Por tanto, para h pequeños,

$$\Delta\mathbf{f}(t; h) \approx d\mathbf{f}(t; h)$$

y

8.2 $$\mathbf{f}(t+h) = \mathbf{f}(t) + \Delta\mathbf{f}(t; h) \approx \mathbf{f}(t) + d\mathbf{f}(t; h).$$

Sea $\mathscr{C}$ la curva descrita por la transformación $\mathbf{f}$ de $[a, b]$. Si $\mathbf{f}'(t) \neq \mathbf{0}$, entonces $d(t; h) = h\mathbf{f}'(t)$ es un vector paralelo al vector tangente a $\mathscr{C}$ en el punto $\mathbf{f}(t)$ (figura 9). La ecuación 8.2 implica que cerca de $\mathbf{f}(t)$ la recta tangente a $\mathscr{C}$ en $\mathbf{f}(t)$ está muy cerca de la curva

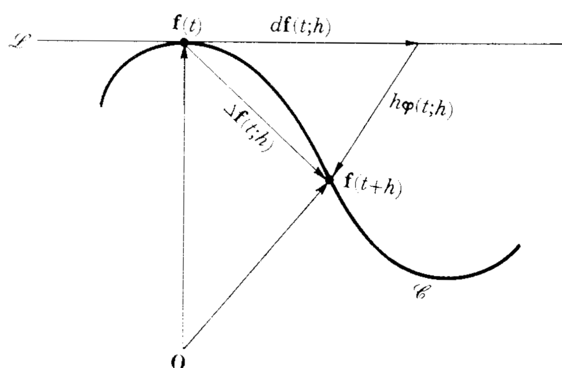


FIGURA 9

Es práctica común usar dt en lugar de h y abreviar $d\mathbf{f}(t; dt)$ por $d\mathbf{f}$. Por tanto,

$$d\mathbf{f} = d\mathbf{f}(t; dt) = \mathbf{f}'(t)dt$$

y $\mathbf{f}'(t)$ es $\dfrac{d\mathbf{f}}{dt}$, una notación ya introducida para la derivada. Cuando usamos $d\mathbf{f}$ para denotar un valor de la diferencial, es generalmente posible determinar por el contexto de la discusión los valores de t y dt que el usuario tiene in mente.

Si $\mathbf{f} = (f_1, \ldots, f_n)$, entonces

$$d\mathbf{f} = \mathbf{f}'(t)dt = (f_1'(t)dt, \ldots, f_n'(t)dt)$$

es decir

8.3 $$d\mathbf{f} = (df_1, \ldots, df_n).$$

Si hacemos $\mathbf{x} = (x_1, \ldots, x_n) = (f_1(t), \ldots, f_n(t)) = \mathbf{f}(t)$, entonces podemos escribir $d\mathbf{x} = d\mathbf{f}$ y $dx_i = df_i$ y, de aquí, 8.3 toma la forma

$$d\mathbf{x} = (dx_1, \ldots, dx_n).$$

De la definición de diferencial y las fórmulas de derivación que hemos desarrollado, se deduce fácilmente que

8.4 $$d(\mathbf{f}+\mathbf{g}) = d\mathbf{f}+d\mathbf{g}$$

8.5 $$d(\mathbf{f}-\mathbf{g}) = d\mathbf{f}-d\mathbf{g}$$

8.6 $$d(\mathbf{f}\cdot\mathbf{g}) = \mathbf{f}\cdot d\mathbf{g}+d\mathbf{f}\cdot\mathbf{g}$$

8.7 $$d(\mathbf{f}\times\mathbf{g}) = \mathbf{f}\times d\mathbf{g}+d\mathbf{f}\times\mathbf{g}$$

8.8 $$d(\varphi\mathbf{f}) = \varphi\, d\mathbf{f}+(d\varphi)\mathbf{f}$$

8.9 $$d(\mathbf{f}\circ\varphi) = (\mathbf{f}'\circ\varphi)d\varphi.$$

Estas fórmulas se verifican bajo las condiciones especificadas en el teorema 7.2 (pág. 123) y en el teorema 7.7 (pág. 126).

La fórmula 8.9 es de especial interés. Si hacemos $\mathbf{x} = \mathbf{f}(t)$ y $t = \varphi(u)$ entonces $\mathbf{x} = \mathbf{f}(\varphi(u)) = \mathbf{g}(u)$, donde $\mathbf{g} = \mathbf{f} \circ \varphi$. En tal caso, la notación $d\mathbf{x}$ para la diferencial parece ambigua; puede querer decir $\mathbf{f}'(t)dt$ o $\mathbf{g}'(u)du$. Sin embargo, esta ambigüedad sólo es aparente, ya que según 8.9

$$\mathbf{g}'(u)du = \mathbf{f}'(t)dt,$$

y, en realidad, es precisamente a causa de esta aparente ambigüedad que la notación diferencial resulta conveniente.

Problemas

1. Determínense $\Delta\mathbf{f}(t; dt)$ y $d\mathbf{f}(t; dt)$, cuando $\mathbf{f}(t) = (t, t^2, t^3)$ y

a) $t = 0,\ dt = 10^{-3}$ *b*) $t = 0,\ dt = 10^{-1}$

c) $t = 0,\ dt = 10^3$ *d*) $t = 10,\ dt = 10^{-1}$

e) $t = 10^3,\ dt = 10^{-1}$ *f*) $t = 10^{10},\ dt = 10^2$

2. Hállese el valor aproximado de $\mathbf{f}(10^{-3})$ cuando

a) $\mathbf{f} = (\cos, \text{sen}, \tan)$

b) $\mathbf{f}(t) = e^{-t}(1, \text{sen}\ t, \cos 2t)$

c) $\mathbf{f}(t) = e^{-t}(1, \text{sen}^2\ t, \cos^2 t)$.

3. Demuéstrese que bajo hipótesis adecuadas

a) $d(\mathbf{f} \cdot \mathbf{g}) = \mathbf{f} \cdot d\mathbf{g} + d\mathbf{f} \cdot \mathbf{g}$

b) $d(\mathbf{f} \circ \varphi) = (\mathbf{f}' \circ \varphi)d\varphi$.

9. INTEGRACIÓN

Una curva puede describirse especificando uno de sus puntos y un vector tangente en cada uno de sus puntos. Supongamos que conocemos que una curva $\mathscr{C}$ pasa por el punto $\mathbf{x}_0$ y que, para cada $t \in [a, b]$, $\mathbf{f}(t)$ es un vector tangente a $\mathscr{C}$. Deseamos determinar una transformación $\mathbf{x}$ de $[a, b]$ tal que $\mathbf{x}(t_0) = \mathbf{x}_0$ para algún $t_0 \in [a, b]$ y $\mathbf{x}'(t) = \mathbf{f}(t)$ para todo $t \in [a, b]$. Entonces $\mathscr{C}$ es descrita por la transformación $\mathbf{x}$ de $[a, b]$.

Para determinar $\mathbf{x}$ debemos resolver la ecuación diferencial

$$\mathbf{x}' = \mathbf{f} \quad \text{sobre } [a, b]$$

sujeta a la condición $\mathbf{x}(t_0) = \mathbf{x}_0$. La solución de esta ecuación diferencial es simple una vez que hemos introducido la integral de una función vectorial.

9.1 Definición. *Si* $\mathbf{f} = (f_1, \ldots, f_n)$ *es una función vectorial definida sobre* $[a, b]$, *entonces*

$$\int_a^b \mathbf{f} = \left(\int_a^b f_1, \ldots, \int_a^b f_n\right).$$

Usamos también la notación $\int_a^b \mathbf{f}(t)\,dt$ para la integral de $\mathbf{f}$ de a a b. Así pues

$$\int_a^b \mathbf{f}\,(t)\,dt = \left(\int_a^b f_1(t)\,dt, \ldots, \int_a^b f_n(t)\,dt\right)$$

La integral $\int_a^b \mathbf{f}$ existe siempre que cada una de las integrales $\int_a^b f_i$, $i = 1, \ldots, n$, existe. En particular, si $\mathbf{f}$ es continua sobre $[a, b]$ entonces $\int_a^b \mathbf{f}$ existe.

El primer teorema fundamental del cálculo —si f es continua sobre un intervalo $\mathscr{I}$ y a, $t \in \mathscr{I}$, entonces $D_t \int_a^t f = f(t)$— puede extenderse a funciones vectoriales como sigue.

9.2 Teorema. *Si* $\mathbf{f} = (f_1, \ldots, f_n)$ *es continua sobre un intervalo* $\mathscr{I}$ *y* $a \in \mathscr{I}$, *entonces*

$$D_t \int_a^t \mathbf{f} = \mathbf{f}(t), \qquad t \in \mathscr{I}.$$

Prueba. La prueba se obtiene por la aplicación del primer teorema fundamental del cálculo a cada una de las funciones componentes:

$$\begin{aligned} D_t \int_a^t \mathbf{f} &= D_t\left(\int_a^t f_1, \ldots, \int_a^t f_n\right) \\ &= \left(D_t \int_a^t f_1, \ldots, D_t \int_a^t f_n\right) \\ &= (f_1(t), \ldots, f_n(t)) \\ &= \mathbf{f}(t). \end{aligned}$$

La extensión del segundo teorema fundamental del cálculo —si F' es continua sobre un intervalo $\mathscr{I}$ y $a, b \in \mathscr{I}$, entonces $\int_a^b F' = F(b) - F(a)$— se obtiene también en forma análoga.

9.3 Teorema. *Si* $\mathbf{F} = (F_1, \ldots, F_n)$ *tiene una derivada continua sobre un intervalo* $\mathscr{J}$, *entonces para todo* $a, b \in \mathscr{J}$

$$\int_a^b \mathbf{F}' = \mathbf{F}(b) - \mathbf{F}(a).$$

PRUEBA. Dejamos la prueba como ejercicio para el estudiante.

Como el teorema fundamental del cálculo puede extenderse a funciones vectoriales, la ecuación diferencial $\mathbf{x}' = \mathbf{f}$ puede resolverse en la forma habitual.

9.4 Teorema. *Si* $\mathbf{f}$ *es continua sobre un intervalo* $\mathscr{J}$, *si* $t_0 \in \mathscr{J}$, *y si* $\mathbf{x}_0$ *es un vector cualquiera, entonces hay una y sólo una solución sobre* $\mathscr{J}$ *de la ecuación diferencial*

$$\mathbf{x}' = \mathbf{f}$$

que satisface $\mathbf{x}(t_0) = \mathbf{x}_0$. *La solución es*

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}.$$

PRUEBA. Supongamos que $\mathbf{x}' = \mathbf{f}$ y $\mathbf{x}(t_0) = \mathbf{x}_0$. Entonces, de acuerdo con el segundo teorema fundamental

$$\int_{t_0}^{t} \mathbf{f} = \int_{t_0}^{t} \mathbf{x}' = \mathbf{x}(t) - \mathbf{x}(t_0)$$

y

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}, \qquad t \in \mathscr{J}.$$

Recíprocamente, si

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}, \qquad t \in \mathscr{J}$$

entonces $\mathbf{x}(t_0) = \mathbf{x}_0$ y según el primer teorema fundamental

$$\mathbf{x}' = \mathbf{f} \quad \text{sobre } \mathscr{J}.$$

Así pues, si una curva $\mathscr{C}$ pasa por el punto $\mathbf{x}_0$ en el tiempo t_0 y $\mathbf{f}(t)$ es un vector tangente a $\mathscr{C}$ para cualquier $t \in [a, b]$, entonces, suponiendo que $\mathbf{f}$ sea continua sobre $[a, b]$, $\mathscr{C}$ está descrita por la transformación $\mathbf{x}$ de $[a, b]$ donde

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}, \qquad t \in [a, b]$$

Sea $\mathbf{x}(t)$ el vector de posición de una partícula P de masa m, $\mathbf{v}(t) = \mathbf{x}'(t)$ la velocidad de P, y $\mathbf{a}(t) = \mathbf{v}'(t)$ la aceleración de P en el instante t. Si la

fuerza que actúa sobre P en el instante t es $\mathbf{F}(t)$, entonces, según la segunda ley del movimiento de Newton, $\mathbf{x}$ debe satisfacer la ecuación

$$m\mathbf{a} = m\mathbf{x}'' = \mathbf{F}.$$

Así pues, la trayectoria de la partícula está determinada por esta ecuación diferencial junto con algunas condiciones iniciales.

9.5 Ejemplo. No teniendo en cuenta la fricción y suponiendo una fuerza gravitacional constante, proporciónese una descripción del movimiento de una partícula de masa m cuya velocidad inicial es $\mathbf{v}_0$ y cuya posición inicial es $\mathbf{x}_0$.

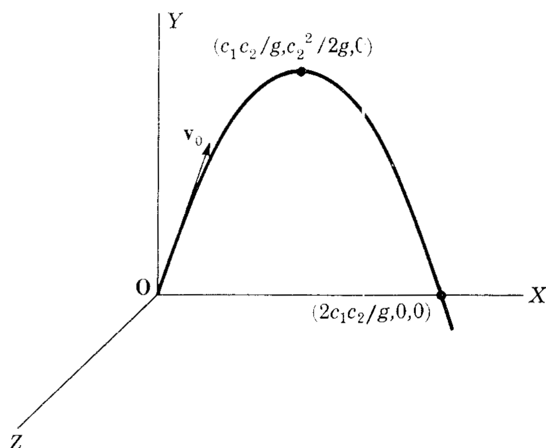


FIGURA 10

SOLUCIÓN. Sea $m\mathbf{g}$ la fuerza constante. Tenemos

$$\mathbf{a} = \mathbf{v}' = \mathbf{g}.$$

Luego

$$\mathbf{v}(t) = \mathbf{v}_0 + \int_0^t \mathbf{g} = \mathbf{v}_0 + \mathbf{g}t$$

y

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_0^t (\mathbf{v}_0 + \mathbf{g}u)\,du$$
$$= \mathbf{x}_0 + \mathbf{v}_0 t + \tfrac{1}{2}\mathbf{g}t^2.$$

Para facilitar el dibujo de la trayectoria de la partícula seleccionaremos un sistema de coordenadas (figura 10) tal que $\mathbf{x}_0 = (0, 0, 0)$, $\mathbf{g} = (0, -g, 0)$ y $\mathbf{v}_0 = (c_1, c_2, 0)$: el origen se coloca en el punto inicial, la fuerza está en la dirección negativa del eje Y, y la dirección del eje X se elige de modo

que $\mathbf{v}_0$ es paralelo al plano XY. Las ecuaciones paramétricas que describen el movimiento de la partícula son, entonces,

$$x = c_1 t$$
$$y = -\tfrac{1}{2}gt^2 + c_2 t$$
$$z = 0.$$

Si $c_1 \neq 0$, éstas son ecuaciones paramétricas de una parábola en el plano XY. La altura máxima de la trayectoria es $\dfrac{c_2^{\,2}}{2g}$ y ésta se alcanza con $t = \dfrac{c_2}{g}$. El eje de la parábola es vertical y su vértice es el punto $\left(\dfrac{c_1 c_2}{g}, \dfrac{c_2^{\,2}}{2g}, 0\right)$.

Problemas

1. Evalúense las siguientes integrales

a) $\displaystyle\int_0^1 (I, I^{1/2}, \exp)$

b) $\displaystyle\int_0^{\pi/2} (\text{sen}\, t, \cos t, \tan t)\, dt$

c) $\displaystyle\int_2^4 \left(\frac{t}{1+t^2}, \sqrt{1+t^2}, 4t^3\right) dt.$

2. Resuélvanse las siguientes ecuaciones diferenciables y dibújese la curva descrita por $\mathbf{x}$ en cada caso.

a) $\mathbf{x}'(t) = \mathbf{c}$, $\mathbf{x}(0) = \mathbf{0}$

b) $\mathbf{x}'(t) = \mathbf{a}t + \mathbf{b}$, $\mathbf{x}(0) = (1, 0, 1)$

c) $\mathbf{x}'(t) = \omega(-\text{sen}\, \omega t, \cos \omega t, 0)$, $\mathbf{x}(0) = (1, C, 0)$.

3. Si no están actuando ningunas fuerzas sobre una partícula de masa m y su posición y velocidad iniciales son $\mathbf{x}_0$ y $\mathbf{v}_0$, respectivamente, descríbase la trayectoria de la partícula.

4. Prescindiendo de los efectos de la atmósfera y suponiendo un suelo perfectamente nivelado estímese la velocidad inicial mínima requerida para hacer que una pelota de golf recorra 250 yardas.

5. ¿Cuál es la contestación al problema 4 si el punto de salida de la pelota está a 25 pies por encima del nivel de la pista ?

6. Puede mostrarse que cada una de las soluciones x de la ecuación diferencial

$$x'' = -\omega^2 x, \quad \omega \text{ una constante}$$

tiene una regla de correspondencia de la forma $x(t) = a\cos(\omega t+\theta)$, $t \in \langle -\infty, \infty \rangle$. Verifíquese que toda función que tiene una regla de correspondencia de esa forma es una solución. Determínese la solución que satisface: *a*) $x(0) = x_0$, $x'(0) = 0$; *b*) $x(0) = 0$, $x'(0) = v_0$, *c*) $x(0) = x_0$, $x'(0) = v_0$

7. ¿Cuál es la forma general de la solución de la ecuación diferencial vectorial

$$m\mathbf{x}'' = -k\mathbf{x}, \quad k > 0, m > 0\,?$$

8. La ecuación diferencial del problema 7 es la ecuación de movimiento de una partícula P de masa m sobre la que actúa una fuerza central que está siempre dirigida hacia $\mathbf{0}$ y cuya magnitud es proporcional a la distancia de la partícula a $\mathbf{0}$.

a) Descríbase el movimiento de la partícula si: 1) $\mathbf{x}(0) = \mathbf{0}$, $\mathbf{x}'(0) = \mathbf{0}$; 2) $\mathbf{x}(0) = \mathbf{x}_0$, $\mathbf{x}'(0) = \mathbf{0}$; 3) $\mathbf{x}(0) = \mathbf{0}$, $\mathbf{x}'(0) = \mathbf{v}_0$.

b) Demuéstrese que la suma de dos soluciones de la ecuación de movimiento es una solución. Descríbase el movimiento de la partícula cuando $\mathbf{x}(0) = \mathbf{x}_0$ y $\mathbf{x}'(0) = \mathbf{v}_0$.

c) Determínense cuáles deben ser la posición y velocidad iniciales de la partícula para que se mueva a lo largo de una circunferencia de radio r alrededor del origen.

10. LONGITUD DE ARCO

Sea $\mathscr{C}$ una curva descrita por la transformación $\mathbf{f}$ de un intervalo cerrado $[a, b]$ en R^n. Consideremos una partición $P = \{t_i \mid i = 0, \ldots, k\}$ de $[a, b]$ donde $a = t_0 < t_1 < \ldots < t_k = b$. Toda partición P de $[a, b]$ define una poligonal constituida por los segmentos rectilíneos de $\mathbf{f}(t_0)$ a $\mathbf{f}(t_1)$, de $\mathbf{f}(t_1)$ a $\mathbf{f}(t_2), \ldots$, de $\mathbf{f}(t_{k-1})$ a $\mathbf{f}(t_k)$. (Esto está ilustrado en la figura 11 para el caso $P = \{t_0, t_1, t_2, t_3, t_4, t_5\}$.) Denotamos la longitud de este arco poligonal por L_P; es decir,

$$L_P = \sum_{i=1}^{k} |\mathbf{f}(t_i)-\mathbf{f}(t_{i-1})|\,.$$

Nuestra idea intuitiva de lo que la longitud de $\mathscr{C}$ será, nos dice que debería sernos posible aproximarnos a la longitud de $\mathscr{C}$ tanto como deséasemos midiendo las longitudes L_P de arcos poligonales como los descritos. Además, como la distancia a lo largo de una línea recta debe ser la distancia más corta entre puntos, L_P debe ser menor que la longitud de $\mathscr{C}$ y, si añadimos puntos a la partición P, la longitud del nuevo arco poligonal debe ser mejor aproximación que la primitiva a la longitud del

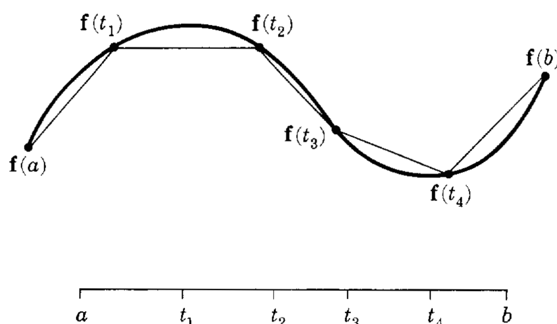


FIGURA 11

arco $\mathscr{C}$. Esto sugiere la siguiente definición. Vamos a denotar en ella por $\mathscr{P}$ al conjunto de todas las particiones del intervalo $[a, b]$.

10.1 Definición. *La curva $\mathscr{C}$ descrita por una transformación* $\mathbf{f}$ *de* $[a, b]$ *se dice que es* ***rectificable*** *si* $\{L_P \mid P \in \mathscr{P}\}$ *tiene una cota superior. Si $\mathscr{C}$ es rectificable,* *la* ***longitud*** *L de $\mathscr{C}$ es el supremo de* $\{L_P \mid P \in \mathscr{P}\}$*; es decir,*

$$L = \sup \{L_P \mid P \in \mathscr{P}\}.$$

Esta longitud de una curva se conforma a las ideas intuitivas antes mencionadas. Como L es una cota superior de $\{L_P \mid P \in \mathscr{P}\}$, L es mayor que o igual a la longitud L_P de cualquier arco poligonal obtenido tomando una partición P de $[a, b]$. Por otra parte, para cualquier $\varepsilon > 0$, existe una partición P de $[a, b]$ tal que $L - \varepsilon < L_P \leqslant L$; de otra forma L no sería el supremo (la cota superior *mínima*) de $\{L_P \mid P \in \mathscr{P}\}$.

Mostramos ahora que si obtenemos una partición P_2 de $[a, b]$ añadiendo algunos puntos a la partición P_1 de $[a, b]$, entonces $L_{P_1} \leqslant L_{P_2}$. A P_2 le llamamos *refinamiento* de P_1.

10.2 Lema. *Si P_2 es un refinamiento de P_1, entonces $L_{P_1} \leqslant L_{P_2}$.*

PRUEBA. Este lema es una simple consecuencia de la desigualdad del triángulo. Sea τ_j el primer punto de P_2 que no está en P_1. Entonces, para algún i, $t_{i-1} < \tau_j < t_i$ y

$$\begin{aligned} |\mathbf{f}(t_i) - \mathbf{f}(t_{i-1})| &= |\mathbf{f}(t_i) - \mathbf{f}(\tau_j) + \mathbf{f}(\tau_j) - \mathbf{f}(t_{i-1})| \\ &\leq |\mathbf{f}(t_i) - \mathbf{f}(\tau_j)| + |\mathbf{f}(\tau_j) - \mathbf{f}(t_{i-1})|. \end{aligned}$$

Por un número finito de tales pasos podemos añadir todos los puntos de P_2 a P_1 y obtener $L_{P_1} \leq L_{P_2}$.

Si tuviésemos que usar la definición para calcular la longitud de una curva, nuestra tarea no sería nada fácil. Sin embargo, para mayoría de

las curvas de interés podemos encontrar la longitud calculando una integral. Consideremos la curva $\mathscr{C}$ descrita por la transformación **f** de $[a, b]$ como la trayectoria de una partícula, donde $\mathbf{f}(t)$ es la posición de la partícula en el instante t. Supongamos que **f** es diferenciable sobre $[a, b]$. Entonces $\mathbf{f}'(t)$ es la velocidad de la partícula en el instante t y $|\mathbf{f}'(t)|$ es la "rapidez" de la partícula en el instante t. Supongamos que tomamos una partición P de $[a, b]$ tal que la velocidad "cambia muy poco" sobre cada arco de $\mathbf{f}(t_{i-1})$ a $\mathbf{f}(t_i)$: digamos que es aproximadamente $\mathbf{f}'(t_i^*)$ sobre este arco donde $t_i^* \in [t_{i-1}, t_i]$. Entonces, usando la noción elemental de que la distancia es igual a la "rapidez" multiplicada por el tiempo, la longitud de $\mathscr{C}$ es aproximadamente

$$S_P = \sum_{i=1}^{k} |\mathbf{f}'(t_i^*)|\,(t_i - t_{i-1}).$$

Reconocemos a S_P como una suma aproximativa de la integral $\int |\mathbf{f}'|$.

Es decir,

$$\int_a^b |\mathbf{f}'| = \lim_{|P| \to 0} S_P$$

donde $|P|$ denota la norma de la partición P:

$$|P| = \text{máx}\,\{t_i - t_{i-1} \mid i = 1, \ldots, k\},$$

y este límite significa:
Para cualquier $\varepsilon > 0$ existe una $\delta > 0$ tal que $|P| < \delta$ implica

$$\left| S_P - \int_a^b |\mathbf{f}'| \right| < \varepsilon.$$

Es, pues, lógico esperar que la longitud de $\mathscr{C}$ sea $\int_a^b |\mathbf{f}|'$.

En la anterior discusión fue necesario que la velocidad "cambiase muy poco" en el arco. Esto quiere decir que necesitamos que $\mathbf{f}'$ sea continua sobre $[a,b]$. Para funciones reales sabemos que si una función g es continua sobre un intervalo cerrado $[a,b]$, entonces g es uniformemente continua sobre $[a,b]$; es decir, para cualquier $\varepsilon > 0$ existe una $\delta > 0$ tal que $|g(x) - g(y)| < \varepsilon$ para todo $x, y \in [a, b]$ con la propiedad de que $|x - y| < \delta$. Definimos la continuidad uniforme para funciones vectoriales en la misma forma y podemos probar que si una función vectorial es continua sobre un intervalo cerrado $[a, b]$ entonces es uniformemente continua sobre $[a, b]$. Probaremos ahora un resultado ligeramente más fuerte.

10.3 Lema. *Si una función vectorial* **g** *es continua sobre un intervalo cerrado* $[a, b]$, *entonces para cada* $\varepsilon > 0$ *hay una* $\delta > 0$ *tal que*

$$|\mathbf{g}(t) - (g_1(s_1), \ldots, g_n(s_n))| < \varepsilon$$

para todo t, $s_1, \ldots, s_n \in [a, b]$ *con la propiedad de que* $|t - s_i| < \delta$ *para* $i = 1, \ldots, n$.

PRUEBA. Como $\mathbf{g}$ es continua sobre $[a, b]$, cada una de las funciones componentes g_i es continua sobre $[a, b]$ y, por tanto, uniformemente continua sobre $[a, b]$. Luego, para $\varepsilon > 0$ hay una $\delta_i > 0$ tal que

$$|g_i(t) - g_i(s_i)| < \frac{\varepsilon}{\sqrt{n}}$$

para todo $t, s_i \in [a, b]$ con la propiedad de que $|t - s_i| < \delta_i$. Haciendo $\delta = \text{mín}\,\{\delta_i \mid i = 1, \ldots, n\}$, tenemos

$$|\mathbf{g}(t) - (g_1(s_1), \ldots, g_n(s_n))| = \left[\sum_{i=1}^{n} |g_i(t) - g_i(s_i)|^2\right]^{1/2} < \left[\sum_{i=1}^{n} \frac{\varepsilon^2}{n}\right]^{1/2} = \varepsilon$$

para t, $s_1, \ldots, s_n \in [a, b]$ con la propiedad de que $|t - s_i < \delta$, $i = 1, \ldots, n$. Y esto completa la prueba.

Ahora estamos en posición de probar la fórmula integral para la longitud de una curva.

10.4 Teorema. *Si* $\mathbf{f}$ *tiene una derivada continua sobre* $[a, b]$, *entonces la curva* $\mathscr{C}$ *descrita por* $\mathbf{f}$ *es rectificable y*

$$L = \int_a^b |\mathbf{f}'|.$$

PRUEBA. En la prueba consideramos $\mathscr{C}$ como una curva en R^3, aunque el método se puede aplicar cualquiera que sea la dimensión del espacio en que la curva esté definida. Sea $P = \{t_0, \ldots, t_k\}$ una partición de $[a, b]$. De acuerdo con el teorema del valor medio (7.9, pág. 126)

$$\begin{aligned} L_P &= \sum_{i=1}^{k} |\mathbf{f}(t_i) - \mathbf{f}(t_{i-1})| \\ &= \sum_{i=1}^{k} |(f_1'(t_i'), f_2'(t_i''), f_3'(t_i'''))| \ (t_i - t_{i-1}) \end{aligned}$$

para algunos $t_i', t_i'', t_i''' \in \langle t_{i-1}, t_i \rangle$. Como f_1', f_2' y f_3' son continuas sobre $[a, b]$, están acotadas sobre $[a, b]$. Supongamos $|f_1'(t)| \leq M_1$, $|f_2'(t)| \leq M_2$ y $|f_3'(t)| \leq M_3$, para toda $t \in [a, b]$. Entonces

$$L_P \leq \sum_{i=1}^{k} \sqrt{M_1^2 + M_2^2 + M_3^2}\,(t_i - t_{i-1}) = (b - a)\sqrt{M_1^2 + M_2^2 + M_3^2}$$

para toda partición P de $[a, b]$. Por tanto, $\{L_P \mid P \in \mathscr{P}\}$ está superiormente acotada y $\mathscr{C}$ es ıectificable; sea L la longitud de $\mathscr{C}$.

Queda por mostrar que $L = \int_a^b |\mathbf{f}'|$. Sea $\varepsilon > 0$ cualquiera. Como $L = \sup \{L_P \mid P \in \mathscr{P}\}$, existe una partición P_1 de $[a, b]$ tal que

10.5 $$L - \varepsilon < L_{P_1} \leq L.$$

Como $|\mathbf{f}'|$ es continua sobre $[a, b]$, $\int_a^b |\mathbf{f}'|$ existe e

$$\int_a^b |\mathbf{f}'| = \lim_{|P| \to 0} S_P = \lim_{|P| \to 0} \sum_{i=1}^{k} |\mathbf{f}'(t_i^*)| \; (t_i - t_{i-1}).$$

Luego existe un $\delta_1 > 0$ tal que

10.6 $$\left| \int_a^b |\mathbf{f}'| - S_P \right| < \varepsilon \quad \text{siempre que } |P| < \delta_1 .$$

Ahora bien

$$\left| \int_a^b |\mathbf{f}'| - L \right| = \left| \int_a^b |\mathbf{f}'| - S_P + S_P - L_P + L_P - L \right|$$
$$\leq \left| \int_a^b |\mathbf{f}'| - S_P \right| + |S_P - L_P| + |L_P - L| .$$

Hemos demostrado que podemos hacer el primero y último términos del segundo miembro de la anterior desigualdad tan pequeños como queramos. Si podemos escoger una partición P tal que la suma de todos los términos del último miembro de la desigualdad sea menor que cualquier número positio dado, habremos entonces demostrado que $L = \int_a^b |\mathbf{f}'|$. De acuerdo con el lema 10.3, existe un $\delta_2 > 0$ tal que

10.7 $$|S_P - L_P| = \left| \sum_{i=1}^{k} |\mathbf{f}'(t_i^*)| \; (t_i - t_{i-1}) - \sum_{i=1}^{k} |(f_1'(t_i'), f_2'(t_i''), f_3'(t_i'''))| \; (t_i - t_{i-1}) \right|$$
$$\leq \sum_{i=1}^{k} |\mathbf{f}'(t_i^*) - (f_1'(t_i'), f_2'(t_i''), f_3'(t_i'''))| \; (t_i - t_{i-1})$$
$$< \varepsilon(b-a) \quad \text{siempre que } |P| < \delta_2 .$$

Luego, si P_2 es un refinamiento de P_1 tal que $|P_2| < \min \{\delta_1, \delta_2\}$,

$$\left| \int_a^b |\mathbf{f}'| - S_{P_2} \right| < \varepsilon \qquad [10.6]$$

$$|S_{P_2} - L_{P_2}| < \varepsilon(b-a) \qquad [10.7]$$

$$|L_{P_2} - L| < \varepsilon \text{ (10.5 y lema 10.2)}.$$

Por tanto,

$$\left| \int_a^b |\mathbf{f}'| - L \right| \leqslant \left| \int_a^b |\mathbf{f}'| - S_{P_2} \right| + |S_{P_2} - L_{P_2}| + |L_{P_2} - L|$$

$$< \varepsilon + \varepsilon(b-a) + \varepsilon.$$

Luego $L = \displaystyle\int_a^b |\mathbf{f}'|$.

10.8 Ejemplo. Sea $\mathscr{C}$ la hélice cilíndrica (figura 4, pág. 23) descrita por $\mathbf{f} = (\cos, \operatorname{sen}, \frac{1}{2}I)$. Determínese la longitud L del arco de $\mathscr{C}$ de $(1, 0, 0)$ a $\left(-1, 0, \dfrac{\pi}{2}\right)$.

Solución. Como $\mathbf{f}(0) = (1, 0, 0)$ y $\mathbf{f}(\pi) = \left(-1, 0, \dfrac{\pi}{2}\right)$,

$$L = \int_0^\pi |\mathbf{f}'| = \int_0^\pi \sqrt{\operatorname{sen}^2 + \cos^2 + \tfrac{1}{4}} = \int_0^\pi \frac{\sqrt{5}}{2} = \frac{\sqrt{5}}{2}\pi.$$

10.9 Ejemplo. Determínese la longitud de la curva $\mathscr{C}$ descrita por $\mathbf{f}(t) = (\cos t, \operatorname{sen} t)$, $t \in [0, 4\pi]$.

Solución. La curva $\mathscr{C}$ es la circunferencia unitaria $\mathscr{C}(\mathbf{0}; 1)$ recorrida dos veces bajo la transformación $\mathbf{f}$ de $[0, 4\pi]$. Usando el teorema 10.4 obtenemos

$$L = \int_0^{4\pi} \sqrt{\operatorname{sen}^2 t + \cos^2 t}\; dt = \int_0^{4\pi} dt = 4\pi.$$

Problemas

1. Determínese la longitud del arco de la parábola descrita por $\mathbf{f}(t) = (t^2, 2t)$, $t \in [0, 1]$.

2. Determínese la longitud de la gráfica de $y = \ln(1-x^2)$ entre $x = 0$ y $x = \frac{1}{2}$.

3. Determínese la longitud de un arco de la cicloide descrita por $\mathbf{f} = a(I - \operatorname{sen}, 1 - \cos)$, donde $a > 0$.

4. Encuéntrese la longitud de la curva descrita por $\mathbf{f}(t) = (t, t, 2t^2)$, $t \in [-3, 3]$.

5. Determínese la longitud del arco de la hélice cónica descrita por $\mathbf{f}(\theta) = (\theta \cos \theta, \theta \operatorname{sen} \theta, \theta)$, $\theta \in [0, 1]$.

6. Determínese la longitud de la curva descrita por la transformación $\mathbf{f}(\varphi) = a\left(\varphi - \operatorname{sen} \varphi, 1 - \cos \varphi, 4 \operatorname{sen} \frac{\varphi}{2}\right)$ del intervalo $[0, 2\pi]$.

7. Considérese la curva $\mathscr{C}$ descrita por

$$x = t$$

$$y = a \cosh \frac{t}{a}$$

$$z = a \operatorname{senh} \frac{t}{a}.$$

Demuéstrese que la distancia a lo largo de $\mathscr{C}$ desde el punto $(0, a, 0)$ hasta un punto $\mathbf{P}_0$ sobre $\mathscr{C}$ es proporcional a la distancia de $\mathbf{P}_0$ al plano XY.

8. Consideremos la elipse descrita por

$$x = a \operatorname{sen} \varphi$$
$$y = b \cos \varphi, \quad \varphi \in [0, 2\pi], a \geq b > 0.$$

Demuéstrese que la longitud de tal elipse es

$$4a \int_0^{\pi/2} \sqrt{1 - e^2 \operatorname{sen}^2 \varphi}\, d\varphi$$

donde $e = \left(1 - \frac{b^2}{a^2}\right)^{1/2}$ es la excentricidad de la elipse. Esta es una integral elíptica de segunda clase. Consúltense tablas y determínese la longitud de la elipse con semieje mayor a y $e = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ y 0.99.

9. La gráfica polar de $r = 1 + \cos \theta$ es una cardioide. Las ecuaciones paramétricas de la cardioide son, por tanto,

$$x = (1 + \cos \theta) \cos \theta$$
$$y = (1 + \cos \theta) \operatorname{sen} \theta, \quad \theta \in [0, 2\pi].$$

Determínese la longitud de la cardioide.

10. Sea g una función real con una derivada continua sobre $[\alpha, \beta]$ y sea $\mathscr{C}$ la gráfica polar de $r = g(\theta)$. Entonces $\mathscr{C}$ está descrita por la función $\mathbf{f} = g\mathbf{u}$ de $[\alpha, \beta]$ donde $\mathbf{u} = (\cos, \operatorname{sen})$. Demuéstrese que la longitud de $\mathscr{C}$ es

$$\int_\alpha^\beta \sqrt{g^2 + (g')^2}$$

11. Resuélvase el problema 9 usando el problema 10.

12. Una ecuación polar de la espiral de Arquímedes es $r = a\theta$. Encuéntrese la longitud de la espiral desde $\theta = 0$ hasta $\theta = 2\pi$.

13. Considérese la parábola cuya ecuación en coordenadas polares es

$$r = \frac{d}{1 - \cos\theta}.$$

Determínese la longitud de la parábola desde el punto sobre el foco hasta el vértice.

14. Si el movimiento de una partícula está descrito por

$$\mathbf{f}(t) = (\cos \omega t, \cos \omega t), \quad \omega > 0,$$

dibújese la trayectoria y encuéntrese la distancia recorrida por la partícula desde el instante $t = 0$ hasta el $t = \dfrac{2\pi}{\omega}$ con y sin integración.

11. TANGENTE UNITARIA, NORMAL PRINCIPAL Y VECTORES BINORMALES

Supongamos que la función $\mathbf{f}$ definida sobre $[a, b]$ tiene una derivada continua distinta de cero sobre $[a, b]$. Entonces la curva $\mathscr{C}$ descrita por la transformación $\mathbf{f}$ de $[a, b]$ se llama *curva lisa*. Como $\mathbf{f}$ tiene una derivada distinta de cero sobre $[a, b]$, la curva $\mathscr{C}$ tiene un vector tangente $\mathbf{f}'(t)$ en cada punto $\mathbf{f}(t)$. Obtenemos el *vector tangente unitario* $\mathbf{T}(t)$ en el punto $\mathbf{f}(t)$ dividiendo el vector tangente $\mathbf{f}'(t)$ por su longitud $|\mathbf{f}'(t)|$, es decir,

11.1
$$\mathbf{T}(t) = \frac{\mathbf{f}'(t)}{|\mathbf{f}'(t)|}.$$

Como la función $\mathbf{f}$ tiene una derivada continua sobre $[a, b]$, la curva $\mathscr{C}$ descrita por $\mathbf{f}$ es rectificable. La longitud $l(t)$ del arco de $\mathscr{C}$ correspondiente la transformación $\mathbf{f}$ de $[a, t]$ es

11.2
$$l(t) = \int_a^t |\mathbf{f}'|, \quad t \in [a, b].$$

El número $l(t)$ es la distancia a lo largo de la curva $\mathscr{C}$ del punto $\mathbf{f}(a)$ al punto $\mathbf{f}(t)$. De acuerdo con el primer teorema fundamental del cálculo, la función l definida por 11.2 tiene una derivada

11.3
$$l' = |\mathbf{f}'|.$$

Luego, usando 11.1, tenemos

11.4 $$\mathbf{f}' = l'\mathbf{T}.$$

Si consideramos la curva lisa $\mathscr{C}$ descrita por **f** como la trayectoria de una partícula, entonces la ecuación 11.4 nos dice que la dirección del vector velocidad $\mathbf{f}'(t)$ es la del vector tangente unitario $\mathbf{T}(t)$ y la magnitud del vector velocidad —la "rapidez"— es $l'(t)$: la razón de cambio de la distancia a lo largo de la curva.

Si $\mathbf{x} = \mathbf{f}(t)$ es la ecuación de una curva en R^3 y si hacemos $s = l(t)$ entonces 11.3 puede escribirse en la forma

11.3′ $$\frac{ds}{dt} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}$$

o, en términos de diferenciales,

$$ds^2 = dx^2 + dy^2 + dz^2.$$

Con esta notación, 11.4 se convierte en

11.4′ $$\frac{d\mathbf{x}}{dt} = \frac{ds}{dt}\mathbf{T}.$$

Supongamos ahora que $\mathbf{f}'$ es diferenciable sobre $[a, b]$; es decir, que $\mathbf{f}''$ existe sobre $[a, b]$. Entonces esgún el problema 3, pág. 127, l'' y $\mathbf{T}'$ existen sobre $[a, b]$ y diferenciando 11.4 obtenemos

11.5 $$\mathbf{f}'' = l''\mathbf{T} + l'\mathbf{T}'.$$

Como $|\mathbf{T}| = 1$ sobre $[a, b]$, sabemos, por el ejemplo 7.3, pág. 124, que $\mathbf{T}'(t)$ es ortogonal al vector tangente $\mathbf{T}(t)$ para todo $t \in [a, b]$.

Cualquier recta que pase por el punto $\mathbf{f}(t)$ de una curva $\mathscr{C}$ y sea ortogonal a la tangente a la curva en ese punto se llama *normal* a la curva. A causa de la significación particular del vector normal $\mathbf{T}'(t)$, la recta que pasa por $\mathbf{f}(t)$ en la dirección de $\mathbf{T}'(t)$ (si $\mathbf{T}'(t) \neq 0$) se llama *normal principal* a la curva $\mathscr{C}$ en $\mathbf{f}(t)$. Si $\mathbf{T}'(t) \neq \mathbf{0}$, entonces definimos el *vector unitario normal principal* $\mathbf{N}(t)$ como sigue:

11.6 $$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|}.$$

Así pues, podemos escribir 11.5 en la forma

11.7 $$\mathbf{f}'' = l''\mathbf{T} + l'|\mathbf{T}'|\mathbf{N},$$

o, lo que es equivalente,

$$\frac{d^2\mathbf{x}}{dt^2} = \frac{d^2s}{dt^2}\mathbf{T} + \frac{ds}{dt}|\mathbf{T}'|\mathbf{N}$$

donde $\mathbf{x} = \mathbf{f}(t)$ y $s = l(t)$.

Si $\mathscr{C}$ es la trayectoria de una partícula que se mueve en R^3, entonces $\mathbf{f}''(t)$ es la aceleración de la partícula en el tiempo t. La ecuación 11.7 nos dice que el vector aceleración se encuentra en el plano determinado por los vectores tangente y normal principal.

11.8 Ejemplo. Determínense los componentes tangencial y normal (normal principal) de $\mathbf{f}''(t)$ en el punto $\mathbf{f}(t)$ de la hélice descrita por $\mathbf{f} = (\cos, \operatorname{sen}, \frac{1}{2}I)$.

SOLUCIÓN 1. De acuerdo con 11.7, el componente tangencial de $\mathbf{f}''(t)$ es $l''(t)$. Como

$$\mathbf{f}(t) = (\cos t, \operatorname{sen} t, \tfrac{1}{2}t),$$

tenemos

$$\mathbf{f}'(t) = (-\operatorname{sen} t, \cos t, \tfrac{1}{2}),$$

$$l'(t) = |\mathbf{f}'(t)| = \tfrac{1}{2}\sqrt{5},$$

y

$$l''(t) = 0.$$

De donde el componente tangencial de $\mathbf{f}''(t)$ es cero, y el componente normal es

$$|\mathbf{f}''(t)| = |(-\cos t, -\operatorname{sen} t, 0)| = 1.$$

SOLUCIÓN 2.

$$\mathbf{T}(t) = \frac{\mathbf{f}'(t)}{|\mathbf{f}'(t)|} = \frac{2}{\sqrt{5}}(-\operatorname{sen} t, \cos t, \tfrac{1}{2})$$

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|} = (-\cos t, -\operatorname{sen} t, 0)$$

y

$$\mathbf{f}''(t) = (-\cos t, -\operatorname{sen} t, 0).$$

De donde,

$$\mathrm{Comp}_{\mathbf{T}(t)}\,\mathbf{f}''(t) = \mathbf{f}''(t)\cdot\mathbf{T}(t) = 0$$

y

$$\mathrm{Comp}_{\mathbf{N}(t)}\,\mathbf{f}''(t) = \mathbf{f}''(t)\cdot\mathbf{N}(t) = 1.$$

En nuestra discusión sobre las curvas hemos definido una curva como una función vectorial continua que tiene un intervalo como dominio. No

hemos hecho ninguna restricción respecto a la dimensión del espacio en el que se encuentra el rango de la función. Así pues, la teoría desarrollada hasta este punto se aplica a una curva en un espacio de dimensión cualquiera. Sin embargo, los ejemplos discutidos nos muestran claramente que nuestro interés principal está en las curvas en R^2 o R^3.

Si $\mathscr{C}$ es una curva en R^2 y $\mathbf{T} = (a, b)$ es un vector tangente unitario a $\mathscr{C}$ en alguno de sus puntos, es fácil ver que el vector unitario normal principal $\mathbf{N}$ en este punto debe ser o $\mathbf{T}^{\perp} = (-b, a)$ o $-\mathbf{T}^{\perp} = (b, -a)$.

Restringimos ahora nuestra atención a las curvas en R^3. El plano que pasa por $\mathbf{f}(t)$ determinado por los vectores $\mathbf{T}(t)$ y $\mathbf{N}(t)$ se llama *plano osculador* de $\mathscr{C}$ en $\mathbf{f}(t)$. Y el vector $\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t)$ se llama *vector binormal*; el binormal es un vector unitario normal al plano osculador. En cada punto $\mathbf{f}(t)$ de $\mathscr{C}$ los vectores $\mathbf{T}(t)$, $\mathbf{N}(t)$ y $\mathbf{B}(t)$ forman un conjunto de vectores unitarios mutuamente ortogonales.

Por ejemplo, en cada punto $\mathbf{f}(t)$ de la hélice descrita por $\mathbf{f} = (\cos, \operatorname{sen}, \frac{1}{2}I)$ (ejemplo 11.8) tenemos

$$\mathbf{T}(t) = \frac{2}{\sqrt{5}}(-\operatorname{sen} t, \cos t, \tfrac{1}{2})$$

$$\mathbf{N}(t) = (-\cos t, -\operatorname{sen} t, 0)$$

$$\mathbf{B}(t) = \frac{1}{\sqrt{5}}(\operatorname{sen} t, -\cos t, 2).$$

Y una ecuación del plano osculador es

$$x \operatorname{sen} t - y \cos t + 2z = t.$$

11.9 Ejemplo. Demuéstrese que si una curva $\mathscr{C}$ se encuentra en el plano $\mathscr{P}$ en R^3 entonces el plano osculador en cualquier punto de $\mathscr{C}$ es $\mathscr{P}$.

SOLUCIÓN. Sea $\mathscr{P} = \{(\mathbf{P} \mid \mathbf{P} \cdot \mathbf{n} = c\}$ y supongamos que $\mathscr{C}$ está descrita por la función $\mathbf{f}$. Como $\mathscr{C} \subset \mathscr{P}$, para cada $t \in \mathscr{D}_{\mathbf{f}}$, $\mathbf{f}(t) \cdot \mathbf{n} = c$. Diferenciando una vez tenemos $\mathbf{f}'(t) \cdot \mathscr{T} = 0$ y, de aquí, $\mathbf{T}(t) \cdot \mathbf{n} = 0$. Diferenciando de nuevo tenemos $\mathbf{T}'(t) \cdot \mathbf{n} = 0$ y, por tanto, $\mathbf{N}(t) \cdot \mathbf{n} = 0$. Esto nos muestra que $\mathbf{n}$ es ortogonal a $\mathbf{T}(t)$ y $\mathbf{N}(t)$. Además, $\mathbf{f}(t)$ pertenece a $\mathscr{P}$ y al plano osculador de $\mathscr{C}$ en $\mathbf{f}(t)$. Por tanto, estos planos deben coincidir.

Problemas

1. Determínense $\mathbf{T}$ y $\mathbf{N}$ para cada una de las siguientes curvas:

a) La parábola: $x = pt^2$, $y = 2pt$.
b) La elipse: $\mathbf{f}(\theta) = (a \cos \theta, b \operatorname{sen} \theta)$, $\theta \in [0, 2\pi]$; $a, b > 0$.
c) La rama de la hipérbola: $x = a \cosh t$, $y = b \operatorname{senh} t$.

d) La hélice cónica: $\mathbf{f}(\theta) = (\theta \cos \theta,\ \theta \operatorname{sen} \theta,\ a\theta)$.
e) La recta: $\mathbf{x} = \mathbf{P}_0 + t\mathbf{a}$.

2. Sea $\mathscr{C}$ la curva descrita por la transformación $\mathbf{f}$ de $[a, b]$. Puede suceder que en un punto $\mathbf{f}(t_0)$ de $\mathscr{C}$ donde $\mathbf{f}'(t_0) = 0$, exista $\lim_{t \to t_0} \mathbf{T}(t)$. En este caso definimos

$$\mathbf{T}(t_0) = \lim_{t \to t_0} \mathbf{T}(t)$$

y a $\mathbf{T}(t_0)$ se le llama *vector unitario tangente* a $\mathscr{C}$ en $\mathbf{f}(t_0)$. Determínese la recta tangente a cada una de las siguientes curvas en los puntos indicados:

a) $\mathbf{f}(t) = (t^3, t^3)$ para $t = 0$
b) $\mathbf{f}(t) = (3t^2, 2+8t^2, -5t^2)$ para $t = 0$
c) $\mathbf{f}(t) = (t^2, t^3, t^4)$ para $t = 0$ y $t = 1$.

3. Cada una de las siguientes es una regla de correspondencia descrita por el movimiento de una partícula (t es el tiempo). En $t = 0$ y $t = 1$ encuéntrense la velocidad, la "rapidez", la aceleración, y las componentes normal y tangencial de la aceleración.

a) $\mathbf{f}(t) = (10 \operatorname{sen} 2\pi t,\ 10 \cos 2\pi t)$
b) $\mathbf{f}(t) = (10 \cos 2\pi t,\ 10 \operatorname{sen} 2\pi t)$
c) $\mathbf{f}(t) = (\cos \pi t^2,\ \operatorname{sen} \pi t^2)$
d) $\mathbf{f}(t) = e^{-t}\left(\cos \frac{\pi}{2} t,\ \operatorname{sen} \frac{\pi}{2} t\right)$
e) $\mathbf{f}(t) = \left(\cos 200\pi t,\ \operatorname{sen} 200\pi t,\ \frac{t}{2\pi}\right)$

4. Sean $r(t)$ y $\theta(t)$ las coordenadas polares de una partícula en el instante t, y sea $\mathbf{x}(t)$ el radio vector que localiza la partícula en el instante t. Entonces

$$\mathbf{x} = r\mathbf{u}$$

donde $\mathbf{u} = (\cos \circ \theta,\ \operatorname{sen} \circ \theta)$; es decir

$$\mathbf{x}(t) = r(t)\,\mathbf{u}(t) = r(t)\,(\cos \theta(t),\ \operatorname{sen} \theta(t))$$

donde $\mathbf{u}(t)$ es un vector unitario radial a la trayectoria (figura 12). Sea $\mathbf{u}^{\perp}(t) = (-\operatorname{sen} \circ \theta,\ \cos \circ \theta)$; $\mathbf{u}^{\perp}(t)$ es $\mathbf{u}(t)$ girado 90° en dirección contraria a la de las manecillas del reloj. Derívense las siguientes fórmulas para la velocidad $\mathbf{v} = \mathbf{x}'$, rapidez $l' = |\mathbf{x}'|$, y aceleración $\mathbf{a} = \mathbf{x}''$ de la partícula:

$$\begin{aligned} \mathbf{v} &= r'\mathbf{u} + r\theta'\mathbf{u}^{\perp} \\ l'^2 &= r'^2 + r^2\theta'^2 \\ \mathbf{a} &= (r'' - r\theta'^2)\mathbf{u} + (r\theta'' + 2r'\theta')\mathbf{u}^{\perp}. \end{aligned}$$

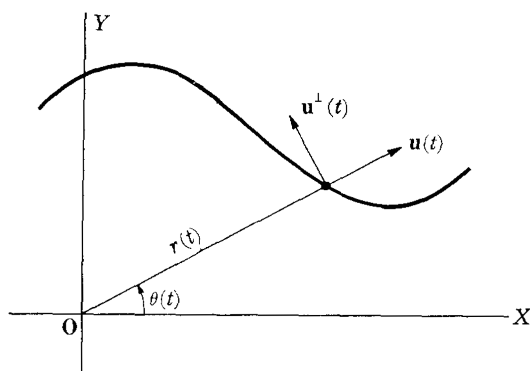


FIGURA 12

5. El movimiento de una partícula se describe en coordenadas polares por:

a) $r(t) = 10,\ \theta(t) = \frac{\pi}{2} - 2\pi t$

b) $r(t) = 10,\ \theta(t) = 2\pi t$

c) $r(t) = 1,\ \theta(t) = \pi t^2$

d) $r(t) = e^{-t},\ \theta(t) = \frac{\pi}{2} t$

e) $r(t) = e^{-t},\ \theta(t) = \frac{\pi}{4} t$

f) $r(t) = \dfrac{4}{1 - \cos \frac{\pi}{2}\left(1 - \frac{t}{2}\right)},\ \theta(t) = \frac{\pi}{2}\left(1 - \frac{t}{2}\right).$

Descríbase en cada caso la trayectoria de la partícula y en $t = 0$ y $t = 1$ determínense la velocidad, la rapidez, la aceleración y las componentes radial, tangencial y normal de la aceleración. [El componente radial es el componente en la dirección de $\mathbf{u} = (\cos \circ \theta, \text{sen} \circ \theta)$.]

6. Si $\mathscr{C}$ es una curva en R^3 descrita por **f** demuéstrese que

$$\mathbf{B} = \frac{\mathbf{f}' \times \mathbf{f}''}{|\mathbf{f}' \times \mathbf{f}''|}, \quad \mathbf{N} = \mathbf{B} \times \mathbf{T} = \frac{(\mathbf{f}' \times \mathbf{f}'') \times \mathbf{f}'}{|(\mathbf{f}' \times \mathbf{f}'') \times \mathbf{f}'|}$$

7. Si una curva está descrita por $\mathbf{f}(t) = (t, t^2, t^3)$, determínense $\mathbf{T}(t)$, $\mathbf{N}(t)$, $\mathbf{B}(t)$ y el plano osculador cuando $t = 0$ y $t = 1$.

8. Determínense **T**, **N** y **B** y el plano osculador en **f**(0) para las curvas en seguida descritas.

a) $\mathbf{f}(t) = (t \cos t, t \operatorname{sen} t, t)$
b) $\mathbf{f}(t) = (t - \operatorname{sen} t, 1 - \cos t, t)$.

9. Si $\mathscr{C}$ es una curva en R^3 y $\mathbf{B}'(t)$ existe, demuéstrese que $\mathbf{B}'(t)$ es paralela a $\mathbf{N}(t)$.

12. CURVATURA Y TORSIÓN

Suponemos en toda esta sección que $\mathscr{C}$ es una curva lisa descrita por una transformación **f** de $[a, b]$. Nuestro objetivo es definir una medida del pandeo de la curva $\mathscr{C}$ en un punto. Sean $\mathbf{f}(t_0)$ y $\mathbf{f}(t_1)$ dos puntos sobre $\mathscr{C}$.

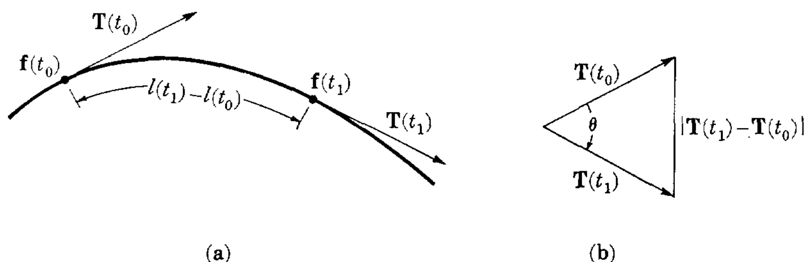


FIGURA 13

Entonces $\mathbf{T}(t_0)$ y $\mathbf{T}(t_1)$ son los vectores unitarios tangentes a $\mathscr{C}$ en los puntos $\mathbf{f}(t_0)$ y $\mathbf{f}(t_1)$ respectivamente (figura 13a). La cantidad $|\mathbf{T}(t_1)-\mathbf{T}(t_0)|$ es una medida de cuánto ha cambiado la dirección de la curva entre $\mathbf{f}(t_0)$ y $\mathbf{f}(t_1)$. En realidad

$$\begin{aligned}|\mathbf{T}(t_1)-\mathbf{T}(t_0)|^2 &= |\mathbf{T}(t_1)|^2 - 2\mathbf{T}(t_0)\cdot\mathbf{T}(t_1) + |\mathbf{T}(t_0)|^2 \\ &= 2(1-\cos\theta) = \left(2 \operatorname{sen}\frac{\theta}{2}\right)^2,\end{aligned}$$

donde θ es el ángulo entre $\mathbf{T}(t_0)$ y $\mathbf{T}(t_1)$ (figura 13b). Nótese que $\left(2 \operatorname{sen}\frac{\theta}{2}\right)^2 \approx \theta^2$ para θ pequeño. Como la longitud del arco de $\mathscr{C}$ desde $\mathbf{f}(t_0)$ hasta $\mathbf{f}(t_1)$ es $|l(t_1)-l(t_0)|$, el cambio promedio de dirección por unidad de distancia sobre este arco es

12.1 $$\frac{|\mathbf{T}(t_1)-\mathbf{T}(t_0)|}{|l(t_1)-l(t_0)|}.$$

Para obtener la razón instantánea del cambio de dirección con respecto a la distancia a lo largo de la curva en el punto $\mathbf{f}(t_0)$, hacemos que t_1 se aproxime a t_0. Si $\mathbf{f}''(t)$ existe, entonces el límite de 12.1 cuando t_1 se aproxima a t_0 también existirá. Este límite es $\dfrac{|\mathbf{T}'(t_0)|}{l'(t_0)}$ y se llama *curvatura* $\kappa(t_0)$ de $\mathscr{C}$ en $\mathbf{f}(t_0)$. Así pues, la curvatura $\kappa(t_0)$ de $\mathscr{C}$ en $\mathbf{f}(t_0)$ se define como

12.2 $$\kappa(t_0) = \frac{|\mathbf{T}'(t_0)|}{l'(t_0)} = \frac{|\mathbf{T}'(t_0)|}{|\mathbf{f}'(t_0)|}.$$

12.3 Ejemplo. Determínese la curvatura de la circunferencia

$$\mathscr{C}(\mathbf{0}; r) = \{(r\cos t, r\operatorname{sen} t) \mid t \in [0, 2\pi]\}.$$

SOLUCIÓN. $\mathscr{C}(\mathbf{0}; r)$ está descrita por la función $\mathbf{f}$, donde

$$\mathbf{f}(t) = (r\cos t, r\operatorname{sen} t).$$

Entonces

$$\mathbf{f}'(t) = (-r\operatorname{sen} t, r\cos t)$$

y

$$\mathbf{f}''(t) = (-r\cos t, -r\operatorname{sen} t).$$

Por tanto

$$|\mathbf{f}'(t)| = r, \quad \mathbf{T}(t) = \frac{1}{r}\mathbf{f}'(t), \quad \mathbf{T}'(t) = \frac{1}{r}\mathbf{f}''(t),$$

y

$$\kappa(t) = \frac{|\mathbf{T}'(t)|}{|\mathbf{f}'(t)|} = \frac{1}{r}.$$

Definimos el *radio de curvatura* $\rho(t)$ de una curva $\mathscr{C}$ en el punto $\mathbf{f}(t)$ como el recíproco de la curvatura en ese punto:

12.4 $$\rho(t) = \frac{1}{\kappa(t)}.$$

En vista del resultado del ejemplo 12.3, el radio de curvatura $\rho(t)$ de $\mathscr{C}$ en $\mathbf{f}(t)$ es el radio de una circunferencia que tiene la curvatura $\kappa(t)$. El punto $\mathbf{f}(t) + \rho(t)\,\mathbf{N}(t)$ se llama *centro de curvatura* de la curva $\mathscr{C}$ correspondiente al punto $\mathbf{f}(t)$, y la circunferencia de radio $\rho(t)$ y centro el centro de curvatura se llama *círculo de curvatura* o *círculo osculador* de $\mathscr{C}$ correspondiente a $\mathbf{f}(t)$.

Como $\kappa = \dfrac{|\mathbf{T}'|}{l'}$, podemos escribir (11.7), pág. 144 como sigue:

12.5 $$\mathbf{f}'' = l''\mathbf{T} + \kappa l'^2\mathbf{N},$$

o, lo que es equivalente,

$$\frac{d^2\mathbf{x}}{dt^2} = \frac{d^2 s}{dt^2}\mathbf{T} + \frac{1}{\rho}\left(\frac{ds}{dt}\right)^2 \mathbf{N}$$

donde $\mathbf{x} = \mathbf{f}(t)$ y $s = l(t)$.

12.6 Ejemplo. Encuéntrese la curvatura de la cúbica alabeada $\mathscr{C}$ descrita por $\mathbf{f}(t) = (t, t^2, t^3)$ en el punto $(1, 1, 1)$.

SOLUCIÓN. El punto $(1, 1, 1)$ es el correspondiente a $t = 1$. Usaremos la fórmula 12.5 para calcular $\kappa(1)$.

$$\mathbf{f}'(t) = (1, 2t, 3t^2), \mathbf{f}'(1) = (1, 2, 3), l'(1) = \sqrt{14}$$
$$\mathbf{f}''(t) = (0, 2, 6t), \mathbf{f}''(1) = (0, 2, 6).$$

Entonces, usando 12.5,

$$l''(1) = \mathbf{f}''(1)\cdot\mathbf{T}(1) = (0, 2, 6)\cdot\frac{1}{\sqrt{14}}(1, 2, 3) = \frac{22}{\sqrt{14}}$$

y

$$\kappa(1)l'^2(1)\mathbf{N}(1) = \mathbf{f}''(1) - l''(1)\mathbf{T}(1) = (0, 2, 6) - \tfrac{22}{14}(1, 2, 3) = \tfrac{1}{7}(-11, -8, 9).$$

Por tanto,

$$\kappa(1) = \tfrac{1}{98}|(-11, -8, 9)| = \tfrac{1}{98}\sqrt{266}.$$

Nota. El ejemplo 12.6 puede resolverse más fácilmente usando la fórmula del problema 2.

12.7 Ejemplo. Si $\mathscr{C}$ es la gráfica de una función g, demuéstrese que

$$\kappa = \frac{|g''|}{[1+(g')^2]^{3/2}}.$$

SOLUCIÓN. Usaremos la fórmula 12.5. $\mathscr{C}$ está descrita por la función $\mathbf{f} = (I, g)$ y tenemos $\mathbf{f}' = (1, g')$ y $\mathbf{f}'' = (0, g'')$. Luego, por 12.5,

$$l'' = \mathbf{f}''\cdot\mathbf{T} = (0, g'')\cdot\frac{(1, g')}{(1+g'^2)^{1/2}} = \frac{g'g''}{(1+g'^2)^{1/2}}$$

y

$$\kappa l'^2\mathbf{N} = \mathbf{f}'' - l''\mathbf{T} = (0, g'') - \frac{g'g''}{(1+g'^2)^{1/2}}\,\frac{(1, g')}{(1+g'^2)^{1/2}}$$

$$= \frac{(-g'g'', g'')}{1+g'^2}.$$

Por tanto,

$$\kappa = \frac{1}{(1+g'^2)^2} |(-g'g'', g'')| = \frac{|g''|}{(1+g'^2)^{3/2}}.$$

Supongamos ahora que $\mathscr{C}$ es una curva en R^3 descrita por $\mathbf{f}$ y que el vector binormal $\mathbf{B}(t)$ es diferenciable en todos los puntos $\mathbf{f}(t)$. En $\mathbf{f}(t)$ el vector $\dfrac{\mathbf{B}'(t)}{l'(t)}$ describe la razón de cambio del vector binormal respecto a la distancia a lo largo de la curva. Como este vector $\dfrac{\mathbf{B}'(t)}{l'(t)}$ es paralelo a $\mathbf{N}(t)$ (problema 9, pág. 149), es igual a un número real por $\mathbf{N}(t)$. El negativo (el inverso aditivo) de este número se llama *torsión* de $\mathscr{C}$ en $\mathbf{f}(t)$ y se denota por $\tau(t)$. Es decir, la torsión τ está definida por la relación

12.8 $$\mathbf{B}' = -\tau l'\mathbf{N}.$$

Como la binormal de una curva plana es constante (ejemplo 11.9, pág. 146, la torsión de una tal curva es cero. Si una curva no es una curva plana, entonces la torsión da una medida del "torcimiento" de la curva respecto al plano osculador.

Por ejemplo, en el caso de la hélice descrita por $\mathbf{f} = (\cos, \operatorname{sen}, \frac{1}{2}I)$, tenemos $\mathbf{N} = (-\cos, -\operatorname{sen}, 0)$, $\mathbf{B} = \dfrac{1}{\sqrt{5}}(\operatorname{sen}, -\cos, 2)$, $\mathbf{B}' = \dfrac{1}{\sqrt{5}}(\cos, \operatorname{sen}, 0)$ y $l' = \frac{1}{2}\sqrt{5}$. Sustituyendo en 12.8, obtenemos

$$\frac{1}{\sqrt{5}}(\cos, \operatorname{sen}, 0) = -\tau\tfrac{1}{2}\sqrt{5}(-\cos, -\operatorname{sen}, 0)$$

de donde, $\tau = \frac{2}{5}$.

Nota. En la mayoría de los casos, la fórmula del problema 8 da un método para obtener τ que es más conveniente que el uso de 12.8.

Problemas

1. Derívese la siguiente fórmula para la curvatura de la curva descrita por $\mathbf{f}$:

$$\kappa = \frac{\sqrt{|\mathbf{f}'|^2 |\mathbf{f}''|^2 - (\mathbf{f}' \cdot \mathbf{f}'')^2}}{|\mathbf{f}'|^3}$$

Sugerencia. Úsese 12.5.

2. Si $\mathscr{C}$ es una curva en R^3 descrita por $\mathbf{f}$, derívese la siguiente fórmula para la curvatura

$$\kappa = \frac{|\mathbf{f}' \times \mathbf{f}''|}{|\mathbf{f}'|^3}$$

3. Determínese la curvatura para cada una de las siguientes curvas:

a) La recta: $\mathbf{x} = \mathbf{P}_0 + t\mathbf{a}$.
b) La parábola: $y = 4px^2$.
c) La elipse: $\mathbf{f}(t) = (a \cos \theta,\ b \operatorname{sen} \theta)$, $\theta \in [0, 2\pi]$; $a, b > 0$.
d) La hélice cilíndrica: $\mathbf{f}(t) = (\cos t,\ \operatorname{sen} t,\ t)$.
e) La hélice cónica: $\mathbf{f}(\theta) = (\theta \cos \theta,\ \theta \operatorname{sen} \theta,\ \theta)$.

4. Sea g una función real con segunda derivada sobre $[\alpha, \beta]$ y sea $\mathscr{C}$ la gráfica polar de $r = g(\theta)$. Derívese la siguiente fórmula para la curvatura de $\mathscr{C}$:

$$\kappa = \frac{|g^2 + 2g'^2 - gg''|}{(g^2 + g'^2)^{3/2}}.$$

5. Determínese la curvatura de las curvas que tienen las siguientes ecuaciones polares:

a) La espiral de Arquímedes: $r = a\theta$.
b) La cardioide: $r = 1 + \cos \theta$.

6. Demuéstrese que

a) $\mathbf{T}' = \kappa l' \mathbf{N}$
b) $\mathbf{N}' = -\kappa l' \mathbf{T} + \tau l' \mathbf{B}$.

Nota. Las dos fórmulas del problema 6 junto con 12.8 se conocen como las fórmulas de Frenet, por el matemático francés F. Frenet. Juegan un importante papel en la geometría de las curvas en el espacio.

7. Si $\mathscr{C}$ es una curva en R^3 descrita por $\mathbf{f}$, úsense 12.5 y el problema 6 para mostrar que

$$\mathbf{f}''' = (l''' - \kappa^2 l'^3)\mathbf{T} + (3\kappa l' l'' + |\kappa|' l'^2)\mathbf{N} + \kappa\tau l'^3 \mathbf{B}.$$

8. Si $\mathscr{C}$ es una curva en R^3 descrita por $\mathbf{f}$, úsense 12.5 y el problema 7 para demostrar que

$$\tau = \frac{(\mathbf{f}' \times \mathbf{f}'') \cdot \mathbf{f}'''}{|\mathbf{f}' \times \mathbf{f}''|^2}.$$

9. Determínese la torsión de la cúbica alabeada descrita por $\mathbf{f}(t) = (t, t^2, t^3)$.

10. Determínese la torsión de la hélice cónica descrita por

$$\mathbf{f}(t) = (t \cos t,\ t \operatorname{sen} t,\ t)$$

en el punto $(0, 0, 0)$.

11. Determínese la torsión de la curva descrita por

$$\mathbf{f}(t) = (t - \operatorname{sen} t,\ 1 - \cos t,\ t)$$

en los puntos correspondientes a $t = 0$, $t = \dfrac{\pi}{2}$, $t = \pi$.

13. APLICACIONES A LA MECÁNICA

Supongamos que $\mathbf{x} = \mathbf{f}(t)$ describe la trayectoria de una partícula en R^3. En la dinámica es común usar el "punto" como notación para la derivada; es, además, práctica general usar "s" en lugar de "l" para la función "longitud de arco". Entonces, la velocidad $\mathbf{v} = \dot{\mathbf{x}}$ viene dada por

13.1 $$\mathbf{v} = \dot{s}\mathbf{T} \qquad [11.4, \text{pág. } 144]$$

y para la aceleración $\mathbf{a} = \dot{\mathbf{v}} = \ddot{\mathbf{x}}$, tenemos la siguiente fórmula:

13.2 $$\mathbf{a} = \ddot{s}\mathbf{T} + \kappa\dot{s}^2\mathbf{N}. \qquad [12.5, \text{pág. } 150]$$

Por 13.1 vemos que la magnitud de la velocidad —la "rapidez"— es $\dot{s}$, la razón de cambio de la longitud de arco a lo largo de la trayectoria, y la dirección de la velocidad es la del vector tangente unitario $\mathbf{T}$. La ecuación 13.2 nos dice que la aceleración se encuentra en un plano determinado por $\mathbf{T}$ y $\mathbf{N}$. Si representamos por $a_\mathrm{T} = \mathbf{a}\cdot\mathbf{T}$, a la *componente tangencial de la aceleración*, y por $a_\mathrm{N} = \mathbf{a}\cdot\mathbf{N}$, a la *componente normal de la aceleración*, tenemos por 13.2

13.3 $$a_\mathrm{T} = \mathbf{a}\cdot\frac{\mathbf{v}}{|\mathbf{v}|} = \ddot{s}$$

y

13.4 $$a_\mathrm{N} = \sqrt{|\mathbf{a}|^2 - a_\mathrm{T}{}^2} = \kappa\dot{s}^2 .$$

Otra expresión para la componente normal de la aceleración puede obtenerse de 13.2 como sigue: como $\mathbf{a}\times\mathbf{T} = a_\mathrm{N}(\mathbf{N}\times\mathbf{T})$ y $|\mathbf{N}\times\mathbf{T}| = 1$,

13.5 $$a_\mathrm{N} = |\mathbf{a}\times\mathbf{T}| = \frac{|\mathbf{a}\times\mathbf{v}|}{|\mathbf{v}|} .$$

13.6 Ejemplo. Si la trayectoria de una partícula está dada por

$$\mathbf{x} = (t^2, \cos t, \operatorname{sen} t),$$

determínense la velocidad, la aceleración y las componentes tangencial y normal de la aceleración.

SOLUCIÓN.

$$\mathbf{v} = \dot{\mathbf{x}} = (2t, -\operatorname{sen} t, \cos t)$$

$$\mathbf{a} = \dot{\mathbf{v}} = (2, -\cos t, -\operatorname{sen} t)$$

$$a_\mathrm{T} = \frac{\mathbf{a}\cdot\mathbf{v}}{|\mathbf{v}|} = \frac{4t + \operatorname{sen} t\cos t - \operatorname{sen} t\cos t}{(4t^2 + \operatorname{sen}^2 t + \cos^2 t)^{1/2}} = \frac{4t}{\sqrt{1+4t^2}}$$

$$a_N = \sqrt{|\mathbf{a}|^2 - a_T{}^2} = \left(4 + \cos^2 t + \operatorname{sen}^2 t - \frac{16t^2}{1+4t^2}\right)^{1/2} = \sqrt{\frac{4t^2+5}{4t^2+1}}.$$

Si una partícula tiene masa m, el vector $m\dot{\mathbf{x}} = m\mathbf{v}$ se llama *momento* (lineal) de la partícula. La segunda ley del movimiento de Newton nos dice: *La razón de cambio del momento es igual a la fuerza*; *es decir*,

13.7 $$D(m\mathbf{v}) = \mathbf{F}.$$

En la mecánica no relativista, m es una constante y esta ecuación de movimiento toma la forma

13.8 $$m\mathbf{a} = \mathbf{F};$$

masa por aceleración es igual a fuerza.

En algunas aplicaciones es conveniente representar la trayectoria de una partícula en forma polar. Consideramos primero el caso especial en que la trayectoria se encuentra en el plano XY y extendemos luego los resultados a trayectorias cualesquiera en R^3. Sean r y θ las coordenadas polares del vector de posición $\mathbf{x}$. Entonces

$$\mathbf{x} = r(\cos\theta, \operatorname{sen}\theta, 0)$$

y

$$\mathbf{v} = \dot{r}(\cos\theta, \operatorname{sen}\theta, 0) + r\dot{\theta}(-\operatorname{sen}\theta, \cos\theta, 0).$$

Si hacemos $\mathbf{u} = (\cos\theta, \operatorname{sen}\theta, 0)$, entonces

$$\mathbf{x} = r\mathbf{u}$$

y

$$\mathbf{v} = \dot{r}\mathbf{u} + r\dot{\mathbf{u}}$$

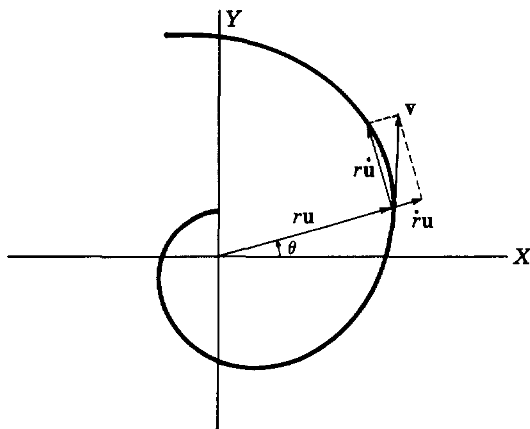


FIGURA 14

donde $\dot{\mathbf{u}} = \dot{\theta}(-\text{sen}\,\theta, \cos\theta, 0)$ es ortogonal a **u** (figura 14). A $\dot{r}$ le llamamos *componente radial de la velocidad* y a $r|\dot{\mathbf{u}}|$ *componente transversa de la velocidad.* El número $|\dot{\mathbf{u}}| = |\dot{\theta}|$ es la razón de cambio del ángulo polar; este número mide la razón angular de giro alrededor del eje Z. Definimos la *velocidad angular* como el vector $\boldsymbol{\omega} = \dot{\theta}\mathbf{k}$. Como

$$\mathbf{u} \times \dot{\mathbf{u}} = (\cos\theta, \text{sen}\,\theta, 0) \times \dot{\theta}(-\text{sen}\,\theta, \cos\theta, 0) = \dot{\theta}\mathbf{k},$$

podemos escribir $\boldsymbol{\omega} = \mathbf{u} \times \dot{\mathbf{u}}$. La velocidad angular es, por tanto, un vector cuya magnitud es la razón de cambio del ángulo polar y cuya dirección es la del eje de rotación y es tal que $\mathbf{u}$, $\dot{\mathbf{u}}$ y $\boldsymbol{\omega}$ forman un sistema levógiro.

Extendemos ahora los anteriores conceptos a cualquier trayectoria $\mathscr{C}$ en $\mathbb{R}^3$. Sea **x** el vector de posición de la partícula y sean

$$r = |\mathbf{x}| \quad \text{y} \quad \mathbf{u} = \frac{\mathbf{x}}{|\mathbf{x}|};$$

r es la distancia de la partícula al origen y **u** es un vector unitario en la dirección del vector de posición **x** (figura 15). Entonces

13.9 $$\mathbf{x} = r\mathbf{u}$$

y a esto se llama ***representación polar*** de $\mathscr{C}$.

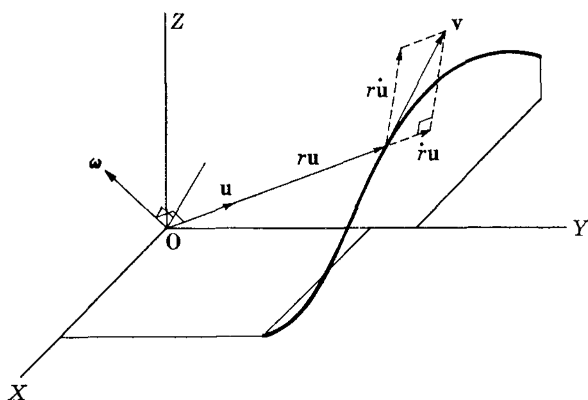


FIGURA 15

Nota. Hemos representado puntos en esta forma polar siempre que hemos usado coordenadas esféricas (pág. 86). En este caso $r = \rho$ y $\mathbf{u} = (\text{sen}\,\varphi \cos\theta, \text{sen}\,\varphi\,\text{sen}\,\theta, \cos\varphi)$.

En la representación polar la velocidad puede escribirse

13.10 $$\mathbf{v} = \dot{r}\mathbf{u} + r\dot{\mathbf{u}}.$$

Como $\mathbf{u}$ es de longitud constante, $\mathbf{u}$ y $\dot{\mathbf{u}}$ son ortogonales. Llamamos a $\dot{r}$ *componente radial de la velocidad* y a $r|\dot{\mathbf{u}}|$ *componente transversa de la velocidad*; y definimos la *velocidad angular* $\boldsymbol{\omega}$ por

13.11 $$\boldsymbol{\omega} = \mathbf{u} \times \dot{\mathbf{u}}.$$

La velocidad angular $\boldsymbol{\omega}$ es un vector en la dirección del eje instantáneo de rotación alrededor del origen y su magnitud $|\dot{\mathbf{u}}|$ es una medida de la razón angular del giro alrededor de este eje.

Usando 13.10 tenemos

13.12 $$\boldsymbol{\omega} = \mathbf{u} \times \dot{\mathbf{u}} = \frac{1}{r}\mathbf{u} \times \mathbf{v} = \frac{\mathbf{x} \times \mathbf{v}}{|\mathbf{x}|^2}.$$

Se sigue del problema 2, pág. 58, que

$$\boldsymbol{\omega} \times \mathbf{u} = (\mathbf{u} \times \dot{\mathbf{u}}) \times \mathbf{u} = -[(\mathbf{u} \cdot \dot{\mathbf{u}})\mathbf{u} - (\mathbf{u} \cdot \mathbf{u})\dot{\mathbf{u}}] = \dot{\mathbf{u}}.$$

Usando esta expresión para $\dot{\mathbf{u}}$ en 13.10 obtenemos

13.13 $$\mathbf{v} = \dot{r}\mathbf{u} + \boldsymbol{\omega} \times \mathbf{x}.$$

Así pues, por ejemplo, si la partícula se mueve con velocidad angular $\boldsymbol{\omega}$ sobre la superficie de una esfera con centro en el origen, entonces $\dot{r} = 0$ y

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{x}.$$

Si $\mathbf{F}$ es una fuerza que actúa en un punto $\mathbf{x}$, el *momento* $\mathbf{L}$ *de la fuerza respecto* $\mathbf{x}_0$ se define por

13.14 $$\mathbf{L} = (\mathbf{x} - \mathbf{x}_0) \times \mathbf{F}.$$

Este vector es perpendicular al plano determinado por $\mathbf{F}$ y $\mathbf{x} - \mathbf{x}_0$ (figura 16) y la magnitud de $\mathbf{L}$ es

$$|\mathbf{L}| = |\mathbf{F}|\,|\mathbf{x} - \mathbf{x}_0| \operatorname{sen} \theta \quad (0 \leq \theta < \pi).$$

Es decir, la magnitud de $\mathbf{L}$ es la magnitud de la fuerza por el brazo de palanca (la distancia de $\mathbf{x}_0$ a la recta de aplicación de la fuerza) y es una medida de la efectividad de $\mathbf{F}$ para producir una rotación alrededor de $\mathbf{x}_0$. El eje de rotación es una recta que pasa por $\mathbf{x}_0$ y es paralela a $\mathbf{L}$ y, si el punto inicial de $\mathbf{L}$ está en $\mathbf{x}_0$ entonces la rotación parece, vista desde la punta de $\mathbf{L}$, como contraria al movimiento de las manecillas del reloj.

Sea $\mathbf{P}$ una partícula de masa m y sea $\mathbf{x}$ el vector de posición de P. El vector $m|\mathbf{x}|^2\boldsymbol{\omega} = \mathbf{x} \times (m\mathbf{v})$ se llama *momento angular* o *momento de la cantidad de movimiento* de P (respecto al origen). Como

$$D[\mathbf{x} \times (m\mathbf{v})] = \mathbf{v} \times (m\mathbf{v}) + \mathbf{x} \times (m\mathbf{a}) = \mathbf{x} \times (m\mathbf{a}),$$

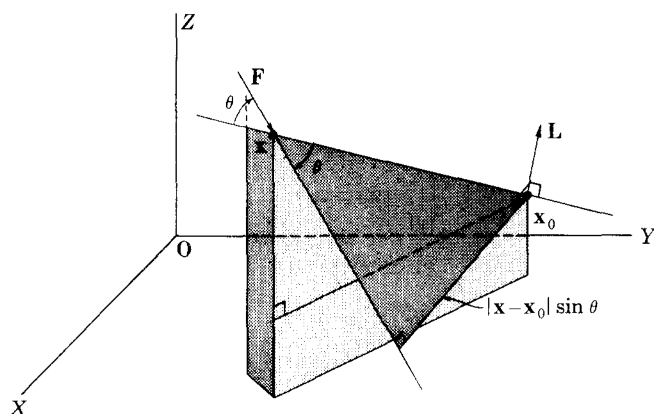


FIGURA 16

la segunda ley de Newton implica

13.15 $$D[\mathbf{x} \times (m\mathbf{v})] = \mathbf{x} \times \mathbf{F} = \mathbf{L}\,;$$

la razón de cambio del momento angular es igual al momento de la fuerza.

13.16 Ejemplo. Una partícula de masa m se mueve sobre una circunferencia de radio r_0 con velocidad angular constante ω_0. Determínese la fuerza que actúa sobre la partícula y el momento angular de la partícula.

SOLUCIÓN. Coloquemos el origen de nuestro sistema de coordenadas en el centro de la circunferencia y orientemos el plano XY de modo tal que: 1) la circunferencia se encuentra en este plano; 2) el movimiento alrededor de la circunferencia visto desde la dirección positiva del eje Z parece contrario al movimiento de las manecillas del reloj, y 3) la partículas cruza el eje X en el instante $t = 0$. Entonces

$$\mathbf{u} = (\cos \omega_0 t,\ \text{sen}\ \omega_0 t,\ 0)$$

y

$$\mathbf{x} = r_0 \mathbf{u}.$$

De donde

$$\mathbf{a} = r_0 \ddot{\mathbf{u}} = -r_0 \omega_0^2 \mathbf{u}$$

y

$$\mathbf{F} = m\mathbf{a} = -m r_0 \omega_0^2 \mathbf{u}.$$

El momento angular es

$$m|\mathbf{x}|^2 \boldsymbol{\omega} = m r_0^2 \omega_0 \mathbf{k}.$$

Nótese también que

$$\dot{s} = |\mathbf{v}| = |r_0 \dot{\mathbf{u}}| = r_0 \omega_0.$$

Expresado en términos de "rapidez" tenemos entonces

$$\mathbf{F} = -\frac{m|\mathbf{v}|^2}{r_0^2}\mathbf{x},$$

$$|\mathbf{F}| = \frac{m|\mathbf{v}|^2}{r_0},$$

$$\boldsymbol{\omega} = \omega_0\mathbf{k} = \frac{|\mathbf{v}|}{r_0}\mathbf{k},$$

y el momento angular es

$$m|\mathbf{x}|^2\boldsymbol{\omega} = mr_0|\mathbf{v}|\mathbf{k}.$$

Problemas

1. La función de posición $\mathbf{x}$ de la partícula de masa m está dada por

a) $\mathbf{x}(t) = 10\left(\cos\frac{2\pi}{T}t,\ \operatorname{sen}\frac{2\pi}{T}t,\ 0\right)$

b) $\mathbf{x}(t) = 10t\left(\cos\frac{2\pi}{T}t,\ \operatorname{sen}\frac{2\pi}{T}t,\ 0\right)$

c) $x(t) = a\operatorname{sen}\frac{2\pi}{T}t,\ y(t) = a\cos\frac{2\pi}{T}t,\ z(t) = b\operatorname{sen}\frac{2\pi}{T}t.$

Determínese en los instantes $t = 0, \frac{T}{2}, T$, la velocidad, la aceleración, las componentes normal y tangencial de la aceleración, la velocidad angular, y el momento angular de la partícula.

2. Demuéstrese que el momento respecto de $\mathbf{x}_0$ de una fuerza $\mathbf{F}$ aplicada en $\mathbf{x}$ no cambia si $\mathbf{F}$ se desliza a lo largo de su recta de acción (la recta por $\mathbf{x}$ paralela a $\mathbf{F}$).

3. Un sistema de dos fuerzas $\mathbf{F}_1$ y $\mathbf{F}_2$ aplicadas en $\mathbf{x}_1$ y $\mathbf{x}_2$ respectivamente se llama *par* si $\mathbf{F}_1+\mathbf{F}_2=\mathbf{0}$. Demuéstrese que la suma $\mathbf{C}$ de los momentos de $\mathbf{F}_1$ y $\mathbf{F}_2$ no depende del punto respecto del cual se calcule el momento; $\mathbf{C}$ se llama *momento del par*. Pruébese, además, que

a) $\mathbf{C} = (\mathbf{x}_1-\mathbf{x}_2)\times\mathbf{F}_1$.

b) La magnitud de $\mathbf{C}$ es la magnitud de $\mathbf{F}_1$ por la distancia entre las rectas de acción de $\mathbf{F}_1$ y $\mathbf{F}_2$.

4. *Una fuerza central respecto a* $\mathbf{0}$ es una que siempre se dirige hacia $\mathbf{0}$ o en la dirección opuesta a $\mathbf{0}$. Demuéstrese que

a) El momento angular respecto a **0** de una partícula sobre la que actúa una fuerza central respecto a **0** es una constante (segunda ley de Kepler del movimiento planetario).
b) La partícula se mueve en un plano que pasa por **0**.

14. RESUMEN

En este capítulo consideramos el cálculo de funciones vectoriales de una variable real. Este material comúnmente se llama cálculo vectorial; el análisis vectorial es álgebra vectorial (capítulo 1) y cálculo vectorial. Vimos que el cálculo de funciones vectoriales de una variable real es análogo y puede en su mayor parte reducirse al cálculo de funciones reales de una variable real. La derivada de una función vectorial puede obtenerse tomando las derivadas de las funciones reales componentes. La integral de una función vectorial se obtiene integrando las funciones reales componentes.

Después de discutir el cálculo vectorial consideramos algunas aplicaciones a la geometría y a la física. La aplicación del cálculo vectorial a la geometría se llama geometría diferencial. En este capítulo presentamos algunos hechos elementales de la geometría diferencial de curvas.

Otros problemas geométricos y físicos exigen el estudio de funciones reales de variable vectorial: funciones con dominio en R_n y rango en R. Estas funciones se estudiarán en el próximo capítulo.

Problemas de repaso

1. Dibújese la curva descrita por **f** cuando

a) $\mathbf{f}(t) = (\cos t, \cos 2t), \mathscr{D}_{\mathbf{f}} = [0, 4\pi]$
b) $\mathbf{f}(t) = (\cos t, \cos t, \cos 2t), \mathscr{D}_{\mathbf{f}} = [0, 4\pi]$.

2. Proporciónese una representación paramétrica de la curva descrita por un punto P sobre una circunferencia de radio uno cuando ésta rueda sobre el lado exterior de una circunferencia de radio 4. Dibújese la curva. Esta curva se llama epicicloide.

3. Si $\mathscr{C}$ tiene la representación paramétrica

$$\mathbf{f}(t) = \left(\cos t, \operatorname{sen} t, \left[\frac{2t}{\pi}\right]\right), \quad t \in [0, 4\pi]$$

determínense todos los puntos en donde $\mathscr{C}$ tiene un vector tangente paralelo a uno de los planos coordenados. Dibújese.

4. Demuéstrese que la hipocicloide definida por las ecuaciones paramétricas

$$x = 4\cos^3\theta$$
$$y = 4\operatorname{sen}^3\theta, \qquad \theta\in\left[-\frac{\pi}{4}, \frac{7\pi}{4}\right]$$

tiene puntos cuspidales en los correspondientes a $\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$.

5. Determínese la longitud de la hipocicloide en el problema 4.

6. Demuéstrese que si una partícula se mueve siempre con "rapidez" constante, su aceleración es siempre ortogonal a su velocidad.

7. ¿Cuándo es cierto que la aceleración y la velocidad de una partícula son paralelas?

8. Encuéntrese la trayectoria $\mathbf{x} = \mathbf{f}(t)$ de una partícula dado que $\mathbf{f}(0) = (0, 0, 1\,600)$, $\mathbf{f}'(t) = (500, 1\,000, -32t)$. ¿Qué distancia recorre la partícula comenzando en el instante $t = 0$ antes de tocar al plano XY? Proporciónense fórmulas para las componentes normal y tangencial de la aceleración. ¿Cuál es el radio de curvatura de la trayectoria cuando $t = 5$?

9. Una partícula se mueve en el plano a lo largo de la espiral $r = e^{\theta}$ con una rapidez constante de 5 pies por segundo.

a) ¿Cuáles son la velocidad y la aceleración de la partícula cuando $\theta = \frac{\pi}{4}$?

b) ¿Cuánto tarda la partícula en ir desde el punto correspondiente a $\theta = 0$ hasta el punto correspondiente a $\theta = \pi$?

c) Si $\theta = 0$ cuando $t = 0$, proporciónense ecuaciones paramétricas para la trayectoria de la partícula.

10. Encuéntrense la tangente unitaria **T**, la normal principal unitaria **N**, la curvatura κ y la longitud **L** de las curvas descritas por

a) $x = a\cos\varphi + a\varphi\operatorname{sen}\varphi$, $y = a\operatorname{sen}\varphi - a\varphi\cos\varphi$, $\quad \varphi\in[0, 2\pi],\ a > 0$

b) $\mathbf{f}(\theta) = (5\cos\theta - \cos 5\theta,\ 5\operatorname{sen}\theta - \operatorname{sen} 5\theta)$.





Funciones reales de un vector

1. INTRODUCCIÓN

En este capítulo discutiremos sobre las funciones con dominio en R^n y rango en R. Una función es una correspondencia de un conjunto de vectores en un conjunto de números reales. Estas funciones también suelen llamarse funciones reales de n variables reales. Los casos donde n es 2 o 3 son los que ocurren con mayor frecuencia en las aplicaciones elementales y son, por consiguiente, los de más interés para nosotros. Sin embargo, como los conceptos fundamentales asociados con funciones de un vector y las propiedades de estas funciones no dependen realmente de la dimensión del espacio (número de las variables), podemos sin añadir dificultad alguna, estudiar el caso general.

Un ejemplo de función real de un vector es la temperatura en un cuarto.

Si establecemos para el cuarto un sistema de coordenadas, definimos la función temperatura T como sigue: en cualquier punto $\mathbf{P} = (x, y, z)$ de la habitación, $T(\mathbf{P})$ es la temperatura en este punto. El dominio de esta función es el conjunto de los puntos en la habitación y el rango es un conjunto de números reales: las temperaturas en los puntos de la habitación. Análogamente, si tenemos en cuenta la dependencia de la temperatura respecto al tiempo, tenemos un ejemplo de una función de cuatro variables reales: $T(x, y, z, t)$ es la temperatura en el punto $\mathbf{P} = (x, y, z)$ en el instante t.

Después de introducir los conceptos de límite, continuidad y derivada para funciones de un vector, presentamos el cálculo diferencial de tales funciones.

En numerosas ocasiones, queremos establecer ciertas restricciones sobre el dominio de una función de un vector. Introducimos ahora terminología para algunos tipos de conjuntos en R^n que tendremos ocasión de usar.

Sea $\mathscr{E}$ un conjunto cualquiera en R^n. El *complemento* de $\mathscr{E}$, representado por $\mathscr{C}\mathscr{E}$, es el conjunto de puntos de R^n que no son de $\mathscr{E}$. Para cualquier punto $\mathbf{x}$ en R^n definimos las siguientes relaciones entre $\mathbf{x}$ y $\mathscr{E}$:

$\mathbf{x}$ es un *punto interior* de $\mathscr{E}$ si existe alguna vecindad de $\mathbf{x}$ que esté contenida en $\mathscr{E}$;
$\mathbf{x}$ es un *punto frontera* de $\mathscr{E}$ si toda vecindad de $\mathbf{x}$ contiene al menos un punto de $\mathscr{E}$ y al menos un punto de $\mathscr{C}\mathscr{E}$;
$\mathbf{x}$ es un punto exterior de $\mathscr{E}$ si hay alguna vecindad de $\mathbf{x}$ que está contenida en $\mathscr{C}\mathscr{E}$.

El conjunto de todos los puntos interiores de $\mathscr{E}$ se llama *interior* de $\mathscr{E}$, y se representa por $\mathscr{E}_i$; el conjunto de puntos frontera de $\mathscr{E}$ se llama *frontera* de $\mathscr{E}$, y se representa por $\mathscr{E}_b$; y el conjunto de todos los puntos exteriores de $\mathscr{E}$ se llama *exterior* de $\mathscr{E}$ y se representa por $\mathscr{E}_e$.

Las siguientes observaciones se deducen en forma evidente de estas definiciones. El interior de $\mathscr{E}$ está contenido en $\mathscr{E}$ y el exterior de $\mathscr{E}$ está contenido en $\mathscr{C}\mathscr{E}$. En realidad, el exterior de $\mathscr{E}$ es el interior de $\mathscr{C}\mathscr{E}$. Un punto frontera de $\mathscr{E}$ puede estar en $\mathscr{E}$ o en $\mathscr{C}\mathscr{E}$. Si llamamos a dos conjuntos *ajenos* cuando tienen una intersección vacía, entonces $\mathscr{E}_i$, $\mathscr{E}_b$ y $\mathscr{E}_e$ son ajenos dos a dos. Además, $R^n = \mathscr{E}_i \cup \mathscr{E}_b \cup \mathscr{E}_e$. Así pues, para cualquier conjunto $\mathscr{E}$ y cualquier punto $\mathbf{x}$ en R^n una y sólo una de las siguientes proposiciones se verifica: $\mathbf{x} \in \mathscr{E}_i$, $\mathbf{x} \in \mathscr{E}_b$, $\mathbf{x} \in \mathscr{E}_e$.

Supongamos que $\mathscr{E} = \{(x, y) \mid y < x\}$. Este es el conjunto de puntos en R^2 que se encuentran debajo de la recta $y = x$ (figura 1). Si $\mathbf{x} = (x, y)$ está en $\mathscr{E}$, entonces $x - y > 0$ y vemos que la vecindad $\mathscr{S}(\mathbf{x}; r)$, donde $r = \dfrac{1}{\sqrt{2}}(x-y)$, se encuentra en $\mathscr{E}$. Así pues, $\mathbf{x}$ es un punto interior de $\mathscr{E}$. Análogamente, podemos demostrar que si $\mathbf{x} = (x, y)$ es un punto por

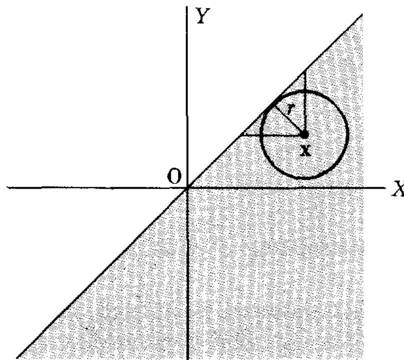


FIGURA 1

encima de la recta $y = x$ (es decir, si $y > x$), entonces $\mathbf{x}$ es un punto exterior de $\mathscr{E}$. Si $\mathbf{x}$ se encuentra en la recta $y = x$, entonces cualquier vecindad de $\mathbf{x}$ contendrá puntos de $\mathscr{E}$ y puntos de $\mathscr{C}\mathscr{E}$ y, por tanto, $\mathbf{x}$ es un punto frontera de $\mathscr{E}$. Así pues, el interior de $\mathscr{E}$ es $\mathscr{E}$, el conjunto de puntos debajo de la recta $y = x$; la frontera de $\mathscr{E}$ es la recta $y = x$; y el exterior de $\mathscr{E}$ es el conjunto de puntos por encima de la recta $y = x$.

Si $\mathscr{E}$ es un conjunto tal que todos los puntos de $\mathscr{E}$ son puntos interiores de $\mathscr{E}$, entonces se dice que $\mathscr{E}$ es *abierto*; es decir, $\mathscr{E}$ es abierto si $\mathscr{E} = \mathscr{E}_i$. Como todo punto de $\mathscr{E}$ es un punto interior o un punto frontera de $\mathscr{E}$, podemos también describir un conjunto abierto como uno que no contiene ninguno de sus puntos frontera. El conjunto $\mathscr{E} = \{(x, y) \mid y < x\}$ que acabamos de estudiar es un ejemplo de un conjunto abierto de R^2. Un conjunto $\mathscr{E}$ se dice que es *cerrado* si su complemento $\mathscr{C}\mathscr{E}$ es abierto; es decir, $\mathscr{E}$ es cerrado si $\mathscr{C}\mathscr{E} = \mathscr{E}_e$ o $\mathscr{E} = \mathscr{E}_i \cup \mathscr{E}_b$. Así pues, un conjunto es cerrado si y sólo si contiene todos sus puntos frontera. Si un conjunto contiene alguno o algunos de sus puntos frontera, pero no todos, entonces no es ni cerrado ni abierto.

1.1 Ejemplo. Demuéstrese que una vecindad $\mathscr{S}(\mathbf{a}; r)$ es un conjunto abierto.

SOLUCIÓN. Si $\mathbf{x} \in \mathscr{S}(\mathbf{a}; r)$, entonces $|\mathbf{x}-\mathbf{a}| < r$. Sea $|\mathbf{x}-\mathbf{a}| = s$. Podemos ver que la vecindad $\mathscr{S}(\mathbf{x}; r-s)$ de $\mathbf{x}$ está contenida en $\mathscr{S}(\mathbf{a}; r)$ como sigue. Si $\mathbf{y} \in \mathscr{S}(\mathbf{x}; r-s)$ entonces

$$|\mathbf{y}-\mathbf{a}| \leqslant |\mathbf{y}-\mathbf{x}|+|\mathbf{x}-\mathbf{a}| < r-s+s = r$$

y, por tanto, $\mathbf{y} \in \mathscr{S}(\mathbf{a}; r)$. Lo que muestra que $\mathscr{S}(\mathbf{a}; r)$ es abierto.

Probamos ahora un par de sencillos resultados sobre conjuntos abiertos.

1.2 Teorema. *Si $\mathscr{E}$ y $\mathscr{F}$ son conjuntos abiertos en* R^n, *entonces* $\mathscr{E} \cap \mathscr{F}$ *es abierto.*

PRUEBA. Si $\mathscr{E} \cap \mathscr{F}$ es vacío, entonces es abierto (problema 3). Supongamos que $\mathscr{E} \cap \mathscr{F}$ no es vacío. Tomemos $\mathbf{x} \in \mathscr{E} \cap \mathscr{F}$. Como $\mathscr{E}$ y $\mathscr{F}$ son abiertos, existen vecindades $\mathscr{S}(\mathbf{x}; r)$ y $\mathscr{S}(\mathbf{x}; s)$ tales que $\mathscr{S}(\mathbf{x}; r) \subset \mathscr{E}$ y $\mathscr{S}(\mathbf{x}; s) \subset \mathscr{F}$. Luego, si $t = \text{mín}\,\{r, s\}$, $\mathscr{S}(\mathbf{x}; t) \subset \mathscr{E} \cap \mathscr{F}$. Lo que muestra que $\mathscr{E} \cap \mathscr{F}$ es abierto.

1.3 Teorema. *Para todo conjunto $\mathscr{E}$ en* R^n, $\mathscr{E}_i$ *es abierto.*

PRUEBA. Si $\mathscr{E}_i$ es vacío, entonces es abierto. Supongamos $\mathscr{E}_i$ no vacío. Si $\mathbf{x} \in \mathscr{E}_i$, entonces hay una vecindad $\mathscr{S}(\mathbf{x}; r)$ de $\mathbf{x}$ que está contenida en $\mathscr{E}$. Como $\mathscr{S}(\mathbf{x}; r)$ es abierto, todo punto $\mathbf{y}$ en $\mathscr{S}(\mathbf{x}; r)$ es un punto interior de $\mathscr{S}(\mathbf{x}; r)$ y, por ello, un punto interior de $\mathscr{E}$. Esto nos muestra que $\mathscr{S}(\mathbf{x}; r)$ está contenida en $\mathscr{E}_i$ y por tanto que $\mathscr{E}_i$ es abierto.

Como $\mathscr{E}_e$ es el interior de $\mathscr{C}\mathscr{E}$, $\mathscr{E}_e$ es abierto. Luego, el complemento de $\mathscr{E}_e$ es un cerrado. El complemento de $\mathscr{E}_e$, que es igual a $\mathscr{E}_i \cup \mathscr{E}_b$ se llama cerradura de $\mathscr{E}$ y se representa por $\bar{\mathscr{E}}$. Recuérdese que ya hemos demostrado que un conjunto $\mathscr{E}$ es cerrado si y sólo si $\mathscr{E} = \mathscr{E}_i \cup \mathscr{E}_b = \bar{\mathscr{E}}$.

Problemas

1. Determínese el interior, la frontera y el exterior de cada uno de los siguientes conjuntos. Dígase si son abiertos, cerrados, o ni abiertos ni cerrados.

a) $\{(x, y) \mid |x-a| < r,\ |y-b| < r\},\ r > 0$
b) $\{(x, y) \mid 4x^2+9y^2 \leqslant 36\}$ *c*) $\{(x, y) \mid 4x^2-9y^2 \leqslant 36\}$
d) $\{(x, y, z) \mid z < x+y\}$ *e*) $\{(x, y, z) \mid x^2+y^2+z^2 \leqslant 4\}$
f) $\{(x, y, z) \mid |x-a| \leqslant r,\ |y-b| < r,\ |z-c| < r\},\ r > 0.$

2. Demuéstrese que los intervalos $\langle a, b\rangle$, $\langle a, \infty\rangle$, $\langle -\infty, b\rangle$, y $\langle -\infty, \infty\rangle$ son conjuntos abiertos en R.

3. Demuéstrese que el conjunto vacío $\varnothing$ y R^n son abiertos y cerrados en R^n.

4. Pruébese que la intersección de un número finito de conjuntos abiertos es un conjunto abierto.

5. Si $\{\mathscr{E}_\alpha \mid \alpha \in \mathscr{J}\}$ es una familia de conjuntos en R^n, la unión de la familia, representada por $\bigcup_{\alpha \in \mathscr{J}} \mathscr{E}_\alpha$, es el conjunto de los puntos $\mathbf{x}$ de R^n tales que $\mathbf{x} \in \mathscr{E}_\alpha$ para algún $\alpha \in \mathscr{J}$. Pruébese que la unión de una familia de conjuntos abiertos es un conjunto abierto.

6. Si $\{\mathscr{E}_\alpha \mid \alpha \in \mathscr{J}\}$ es una familia de conjuntos en R^n, la intersección de

la familia, representada por $\bigcap_{\alpha \in \mathscr{I}} \mathscr{E}_\alpha$, es el conjunto de puntos $\mathbf{x}$ de R^n tales que $\mathbf{x} \in \mathscr{E}_\alpha$ para todo $\alpha \in \mathscr{I}$. Muéstrese que la intersección de una familia de conjuntos abiertos no es necesariamente abierta.

7. Usando la terminología introducida en los problemas 5 y 6 muéstrese que

a) $\mathscr{C} \bigcup_{\alpha \in \mathscr{I}} \mathscr{E}_\alpha = \bigcap_{\alpha \in \mathscr{I}} \mathscr{C}\mathscr{E}_\alpha$ y $\mathscr{C} \bigcap_{\alpha \in \mathscr{I}} \mathscr{E}_\alpha = \bigcup_{\alpha \in \mathscr{I}} \mathscr{C}\mathscr{E}_\alpha$.

b) la intersección de una familia de conjuntos cerrados es un conjunto cerrado.

c) la unión de una familia finita de conjuntos cerrados es cerrada.

8. Si $\mathscr{E}$ es un conjunto cualquiera en R^n demuéstrese que $\mathscr{E}_b$ es cerrado.

9. Si $\mathscr{E}$ es un conjunto cualquiera en R^n demuéstrese que

a) $\mathscr{E}_i$ es el mayor conjunto abierto contenido en $\mathscr{E}$.

b) $\bar{\mathscr{E}}$ es el menor conjunto cerrado que contiene a $\mathscr{E}$.

2. FUNCIONES REALES DE UN VECTOR: GRÁFICAS

Una *función real f de un vector* es una correspondencia de un conjunto $\mathscr{A}$ de vectores a un conjunto $\mathscr{B}$ de números reales tal que para cada $\mathbf{a} \in \mathscr{A}$ existe un y sólo un elemento $f(\mathbf{a}) \in \mathscr{B}$ que denominamos su correspondiente. Así pues, una tal función puede considerarse como una transformación del conjunto $\mathscr{A}$ en R^n sobre un conjunto de números reales. Sugiere esto una terminología conveniente para estas funciones: funciones de R^n en R.

Por ejemplo, la función f con dominio R^2 y regla de correspondencia $f(\mathbf{x}) = |\mathbf{x}|$, donde $\mathbf{x} \in R^2$, transforma el espacio R^2 sobre el intervalo $[0, \infty\rangle$; cada punto de R^2 se transforma en el número real que es la distancia del punto al origen. Puede darse una imagen geométrica de la función distinguiendo los puntos que se transforman sobre el mismo número. Queremos, pues, determinar cuál es el conjunto de los puntos que se transforman sobre el número c, donde $c \geqslant 0$. Este es el conjunto $\{\mathbf{x} \mid |\mathbf{x}| = c\}$. Si $c = 0$, el conjunto se compone del punto único $\mathbf{0}$; si $c > 0$, el conjunto es una circunferencia con centro en $\mathbf{0}$ y radio c. Marcando como c al conjunto $\{\mathbf{x} \mid f(\mathbf{x}) = c\}$, obtenemos la deseada imagen geométrica de la función (figura 2).

En general, una función f de R^2 a R puede representarse geométricamente localizando los conjuntos de puntos donde la función tiene el mismo valor y marcándolos con tal valor. Estos conjuntos se llaman *curvas de nivel* de la función. La curva de nivel correspondiente a un número c es el conjunto $\{(x, y) \mid f(x, y) = c\}$.

Los mapas de contornos o topográficos son ejemplos de este medio de representar una función. Las desigualdades del terreno se muestran en el mapa trazando las curvas de nivel —todos los puntos que están a la misma altura. En los mapas meteorológicos se dibujan las curvas de igual presión barométrica. Estas curvas se llaman isobaras. Y en física, la relación entre la presión y el volumen de un gas ideal ($T = pv$) con la temperatura mantenida constante está representada geométricamente por tal mapa. En este caso las curvas de nivel se llaman curvas isotermas.

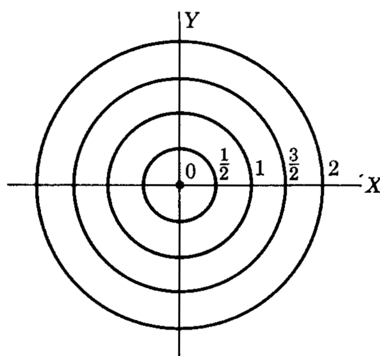


FIGURA 2

Si f es una función de R^2 en R, la *gráfica* de f es el conjunto de puntos $\{(x, y, z) \mid z = f(x, y), (x, y) \in \mathscr{D}_f\}$. Este conjunto es una superficie en R^3, y un dibujo en perspectiva de la gráfica es otro medio de obtener una representación geométrica de una función de R^2 en R.

Consideremos de nuevo la función f con dominio R^2 y regla de correspondencia $f(\mathbf{x}) = |\mathbf{x}|$ o $f(x, y) = \sqrt{x^2 + y^2}$. Al dibujar la gráfica de f usamos la información ganada en la anterior discusión sobre las curvas de nivel

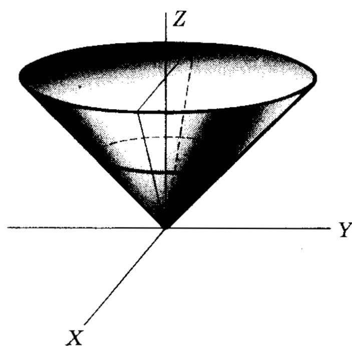


FIGURA 3

de esta función. No hay puntos de la gráfica bajo el plano XY, ya que $z = \sqrt{x^2+y^2} \geqslant 0$. En el plano XY ($z = 0$) el único punto de la gráfica es el $\mathbf{0}$. El conjunto de puntos en la gráfica de f que se encuentran en el plano $z = c > 0$, es una circunferencia de radio c. En general, las curvas de nivel de la función son las curvas de intersección de la gráfica con los planos paralelos al plano XY. Análogamente, ayuda al dibujo de la gráfica localizar la intersección de la gráfica con los planos XZ y YZ. La intersección de la superficie de la ecuación $z = \sqrt{x^2+y^2}$ con el plano XZ ($y = 0$) viene dada por la ecuación $z = \sqrt{x^2} = |x|$ y la intersección con el plano YZ ($x = 0$) viene dada por $z = \sqrt{y^2} = |y|$. La gráfica aparece dibujada en la figura 3. Es un cono generado por la rotación alrededor del eje Z del rayo $z = x$, $x \geqslant 0$, en el plano XZ.

2.1 Ejemplo. Proporciónese una representación geométrica de la función f de dominio R^2 y regla de correspondencia $f(x, y) = x^2+2xy$.

SOLUCIÓN. La curva de nivel sobre la que la función tiene el valor c es el conjunto $\{(x, y) \mid x^2+2xy = c\}$. Si $c = 0$, entonces la curva de nivel consiste en las rectas $x = 0$ y $y = -\frac{1}{2}x$. Si $c \neq 0$, la curva de nivel es la hipérbola $y = \dfrac{c}{2x} - \dfrac{x}{2}$ (figura 4a). Esta hipérbola tiene asíntotas $x = 0$ y $y = -\frac{1}{2}x$.

La gráfica de f es el conjunto $\{(x, y, z) \mid z = x^2+2xy\}$. La intersección de esta superficie con el plano $y = mx$ está dada por $z = (1+2m)x^2$; esta es una parábola si $m \neq -\frac{1}{2}$. Un dibujo de la gráfica de f se da en la figura 4b.

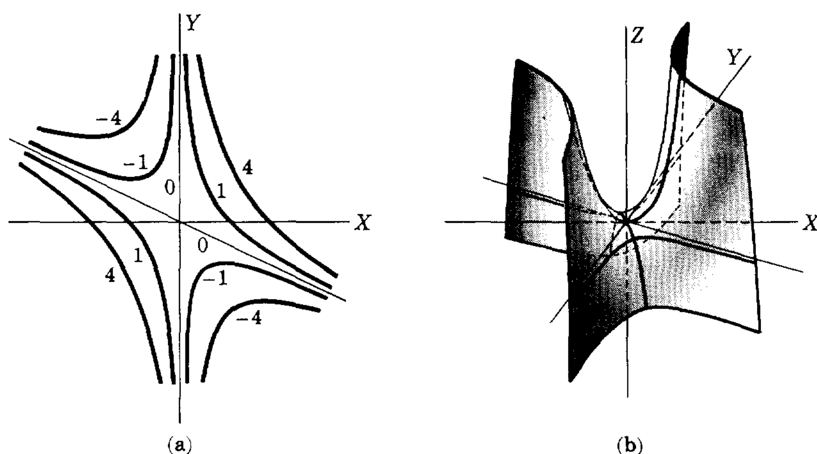


FIGURA 4

Consideremos ahora la gráfica de una función f de R^3 en R. Es ésta el conjunto $\{(x, y, z, w) \mid w = f(x, y, z)\}$, un conjunto de R^4. Aunque la gráfica es aún un concepto útil, no podemos visualizar un conjunto en un espacio de cuatro dimensiones. Sin embargo, se obtiene una imagen gráfica de una función f de R^3 en R determinando el conjunto de puntos que se transforman en un número especificado c; es decir, el conjunto $\{(x, y, z) \mid f(x, y, z) = c\}$. A este conjunto se le llama *superficie de nivel* de f.

Un ejemplo físico de una función de R^3 en R es la función potencial de un campo eléctrico. En este caso las superficies de nivel de la función se llaman superficies equipotenciales.

2.2 Ejemplo. Sea f la función con dominio R^3 y regla de correspondencia $f(\mathbf{x}) = |\mathbf{x}|$. Descríbanse las superficies de nivel de esta función.

SOLUCIÓN. Las superficies de nivel de esta función son los conjuntos $\{\mathbf{x} \mid |\mathbf{x}| = c\}$. Si $c < 0$, el conjunto es vacío; si $c = 0$, el conjunto es el punto único $\mathbf{0}$; si $c > 0$, el conjunto es una superficie esférica con centro $\mathbf{0}$ y radio c (figura 5).

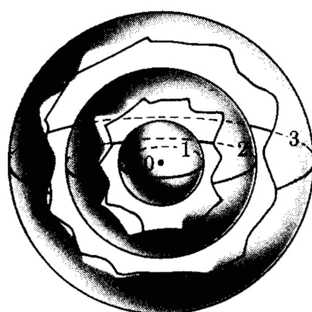


FIGURA 5

Nota. Frecuentemente describimos una función enunciando tan solo una regla de correspondencia. Ha de entenderse entonces que el dominio de la función consiste en todos los puntos sobre los que la regla puede aplicarse con sentido.

Problemas

1. Determínense las curvas de nivel y dibújense las gráficas de cada una de las funciones siguientes:

a) $f(x, y) = x + y^2$ *b*) $f(x, y) = 2x - y$

c) $f(x, y) = x^2 + y^2$ *d*) $f(x, y) = 2x^2 + y^2$

e) $f(x, y) = 2x^2 - y^2$ *f*) $f(x, y) = \dfrac{x}{x+y}$

g) $f(x, y) = x^3 - y$ *h*) $f(x, y) = e^{xy}$

i) $f(x, y) = \operatorname{sen}(x+y)$ *j*) $f(x, y) = \sqrt{x} - y$.

2. Dibújense las superficies de nivel correspondientes a $c = -2, 0, 2$ para las siguientes funciones:

a) $f(x, y, z) = x^2 + y^2 - z$ *b*) $f(x, y, z) = z - \sqrt{x^2 + y^2}$

c) $f(x, y, z) = x^2 + y^2$ *d*) $f(x, y, z) = 3x + 4y - z$.

3. Si $f^*(\mathscr{E}) = \{\mathbf{x} \mid f(\mathbf{x}) \in \mathscr{E}\}$, demuéstrese que

a) $f^*(\mathscr{E} \cup \mathscr{F}) = f^*(\mathscr{E}) \cup f^*(\mathscr{F})$ *b*) $f^*(\mathscr{E} \cap \mathscr{F}) = f^*(\mathscr{E}) \cap f^*(\mathscr{F})$.

3. OPERACIONES SOBRE FUNCIONES

En el estudio de funciones reales de variable real, comenzamos con ciertas funciones básicas. Estudiamos sus propiedades, vimos cuán grandes clases de funciones podían formarse por combinación de las funciones básicas, y cómo podían derivarse propiedades para las funciones más complejas en base a las que sabíamos tenían las funciones básicas. Seguimos el mismo procedimiento en nuestro estudio de las funciones reales de un vector. Ejemplos de funciones básicas entre las funciones reales de una variable real son las funciones constantes y la función identidad. Aquí, las funciones correspondientes son las funciones constantes y las funciones proyección.

3.1 Definición. *La función constante c es la función con* R^n *como dominio y cuyo rango consiste en tan solo el número c.*

3.2 Definición. *La función proyección* I_k $(k = 1, \ldots, n)$ *es la función con* R^n *como dominio y con regla de correspondencia* $I_k(\mathbf{x}) = x_k$, *donde* x_k *es el componente k-ésimo de* $\mathbf{x}$.

Así pues, la función constante c transforma el espacio total R^n sobre el número, único, c. La función de proyección I_k transforma cada punto de R^n sobre su proyección ortogonal sobre el eje X_k; es decir, $I_1(x, y, z) = x$, $I_2(x, y, z) = y$, e $I_3(x, y, z) = z$.

Para las funciones de R^n en R definimos las mismas operaciones que las que definimos para funciones de R en R: adición, sustracción, multiplicación, división y composición.

3.3 Definición. *Si f y g son funciones de* R^n *en* R *con dominios respectivos* $\mathscr{D}_f$ *y* $\mathscr{D}_g$, *entonces* $f+g$, $f-g$, fg, *y* f/g *se definen como sigue:*

$$\mathscr{D}_{f+g} = \mathscr{D}_f \cap \mathscr{D}_g, \qquad [f+g](\mathbf{x}) = f(\mathbf{x})+g(\mathbf{x})$$

$$\mathscr{D}_{f-g} = \mathscr{D}_f \cap \mathscr{D}_g, \qquad [f-g](\mathbf{x}) = f(\mathbf{x})-g(\mathbf{x})$$

$$\mathscr{D}_{fg} = \mathscr{D}_f \cap \mathscr{D}_g, \qquad [fg](\mathbf{x}) = f(\mathbf{x})\,g(\mathbf{x})$$

$$\mathscr{D}_{f/g} = \{\mathbf{x}\in\mathscr{D}_f \cap \mathscr{D}_g \mid g(\mathbf{x}) \neq 0\}, \qquad \left[\frac{f}{g}\right](\mathbf{x}) = \frac{f(\mathbf{x})}{g(\mathbf{x})}.$$

Es fácil ver por las anteriores definiciones que la adición y la multiplicación de funciones de R^n en R son operaciones asociativa y conmutativa; también se verifica la ley distributiva:

$$f+(g+h) = (f+g)+h \qquad f(gh) = (fg)h$$
$$f+g = g+f \qquad fg = gf$$
$$f(g+h) = fg+fh.$$

3.4 Definición. *Si f es una función de* R^n *en* R *y g es una función de* R *en* R, *entonces la composición de g con f, que representamos por $g \circ f$, se define como sigue*

$$\mathscr{D}_{g\circ f} = \{\mathbf{x} \mid \mathbf{x}\in\mathscr{D}_f,\ f(\mathbf{x})\in\mathscr{D}_g\}, \quad [g\circ f](\mathbf{x}) = g(f(\mathbf{x})).$$

Todas las funciones dadas en la sección previa pueden expresarse como combinaciones de funciones de R en R y las funciones básicas definidas en 3.1 y 3.2. Por ejemplo,

si $f(x, y) = x^2+2xy, \qquad f = I_1{}^2+2I_1I_2$

si $f(x, y) = \dfrac{x}{x+y}, \qquad f = \dfrac{I_1}{I_1+I_2}$

si $f(x, y) = e^{xy}, \qquad f = \exp\circ(I_1I_2)$

si $f(x, y, z) = z-\sqrt{x^2+y^2}, \qquad f = I_3-I^{1/2}\circ(I_1{}^2+I_2{}^2).$

Una *función polinomial* de R^n en R es una función que puede obtenerse de las funciones constantes y proyección efectuando las operaciones de adición, sustracción y multiplicación un número finito de veces. Por ejemplo,

$$p = 3I_1{}^2I_2I_3-4I_1{}^4I_3{}^2+10I_3{}^2$$

es una función polinomial de R^3 en R. La regla de correspondencia para esta función es

$$p(x, y, z) = 3x^2yz-4x^4z^2+10z^2$$

y

$$\mathscr{D}_p = \mathrm{R}^3.$$

Una *función racional* r de R^n en R es un cociente de polinomiales, es decir, $r = p/q$ donde p y q son polinomiales de R^n en R. Por ejemplo,

$$r = \frac{3I_1I_2 + I_2^3}{I_1^2 + I_1I_2^2}$$

es una función racional de R^2 en R. La regla de correspondencia para esta función es

$$r(x, y) = \frac{3xy + y^3}{x^2 + xy^2}$$

y

$$\mathscr{D}_r = \{(x, y) \mid x^2 + xy^2 \neq 0\}.$$

Problemas

1. Proporciónese el dominio y regla de correspondencia para cada una de las siguientes funciones de R^3 en R.

a) $f = I_1I_2^2 + I_3$ *b*) $f = 2I_2I_3 + I_1^3$

c) $f = \dfrac{I_1I_2I_3}{I_1 + I_3}$ *d*) $f = \dfrac{I_1^2 + 3I_3}{I_1 - I_2}$.

2. Si $f = 2I_1 + I_2$ y $g = \dfrac{I_1I_2}{I_1^2 + I_2^2}$, determínense $f+g$, $f-g$, fg, y f/g en el punto (2, 1).

3. Si $f = I_1I_2$ y $g = \ln$, determínese $g \circ f$ en (2, 1).

4. Proporciónese el dominio y regla de correspondencia de cada una de estas funciones de R^3 en R.

a) $f = I^{1/2} \circ (I_3 + \cos \circ I_1I_2)$ *b*) $f = \ln \circ (2I_1 + I_3)$

c) $f = \exp \circ (I_1I_2I_3)$.

5. Proporciónese f como una combinación de funciones cuando

a) $f(x, y) = xy^2 + y$

b) $f(x, y) = \text{sen}\,(x^2 + y^2)$.

6. Si f es una función de R^n en R y g y h son funciones de R en R, pruébese que:

a) $h \circ (g \circ f) = (h \circ g) \circ f$ *b*) $(g+h) \circ f = g \circ f + h \circ f$

c) $(gh) \circ f = (g \circ f)(h \circ f)$.

4. LÍMITES

En esta sección discutiremos el concepto de límite de una función real de un vector. Por nuestra experiencia anterior con límites, esperamos que el "límite de f en $\mathbf{a}$ es b" (lo que escribiremos: $\lim_{\mathbf{a}} f = b$ o $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = b$) significará que $f(\mathbf{x})$ está próximo a b cuando $\mathbf{x}$ está próximo a $\mathbf{a}$. Definimos, pues, el límite de f solamente en puntos de acumulación de su dominio. El punto $\mathbf{a}$ es un punto de acumulación de $\mathscr{D}_f$ si toda vecindad reducida de $\mathbf{a}$, $\mathscr{S}'(\mathbf{a};\, \delta)$, contiene un punto de $\mathscr{D}_f$.

Damos ahora una definición analítica de $\lim_{\mathbf{a}} f = b$, suponiendo que $\mathbf{a}$ es un punto de acumulación de $\mathscr{D}_f$.

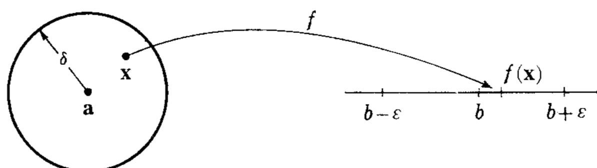


FIGURA 6

4.1 Definición. *El número b se dice que es el* ***límite de la función f en $\mathbf{a}$*** *si para cada número $\varepsilon > 0$, hay un número $\delta > 0$, tal que*

$$|f(\mathbf{x})-b| < \varepsilon$$

siempre que $\mathbf{x}\in\mathscr{D}_f$ y $0 < |\mathbf{x}-\mathbf{a}| < \delta$.

Geométricamente la definición nos dice que $\lim_{\mathbf{a}} f = b$ si para cualquier vecindad dada $\mathscr{S}(b;\, \varepsilon)$ de b existe una vecindad $\mathscr{S}(\mathbf{a};\, \delta)$ de $\mathbf{a}$ tal que $\mathbf{x}$ en $\mathscr{S}'(\mathbf{a};\, \delta)$ y $\mathbf{x}$ en $\mathscr{D}_f$ implica que $f(\mathbf{x})$ está en $\mathscr{S}(b;\, \varepsilon)$; es decir, si f transforma $\mathscr{D}_f \cap \mathscr{S}'(\mathbf{a};\, \delta)$ en $\mathscr{S}(b;\, \varepsilon) = \langle b-\varepsilon, b+\varepsilon\rangle$ (figura 6).

Damos ahora un ejemplo para ilustrar el concepto de límite de una función real de un vector.

4.2 Ejemplo. Si f es la función definida por $f(x, y) = x^2+2xy$, encuéntrese $\lim_{(x,y)\to(3,\,-1)} f(x, y)$ si es que existe.

SOLUCIÓN. (Figura 4, pág. 169.) Si (x, y) está próximo a $(3, -1)$, entonces $f(x, y)$ está próximo a 3. Esperamos, pues, que el límite de f en $(3, -1)$ sea 3. Para verificarlo, para cada $\varepsilon > 0$ debemos encontrar una $\delta > 0$ tal que

$$|x^2+2xy-3| < \varepsilon$$

siempre que $0 < |(x, y)-(3, -1)| < \delta$.

En la figura 7, la región sombreada se transforma en $\langle 3-\varepsilon, 3+\varepsilon\rangle$. Nuestro propósito es encontrar una $\delta > 0$ tal que $\mathscr{S}'((3, -1); \delta)$ se encuentre en la región sombreada. No intentamos determinar la mayor de tales δ sino que nos contentamos con encontrar de un modo sencillo una δ adecuada.

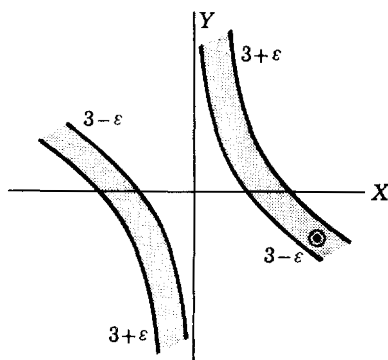


FIGURA 7

Expresamos $|x^2+2xy-3|$ en términos de $|x-3|$ y $|y+1|$. Como $|x-3| \leqslant |(x, y)-(3, -1)|$ y $|y+1| \leqslant |(x, y)-(3, -1)|$, la magnitud de estos términos puede controlarse con la elección de δ.

$$\begin{aligned}|x^2+2xy-3| &= |(x-3)^2+2(x-3)(y+1)+4(x-3)+6(y+1)| \\ &\leqslant |x-3|^2+2|x-3|\,|y+1|+4|x-3|+6|y+1|.\end{aligned}$$

Para simplificar esta expresión podemos restringir la elección de δ de modo que $\delta \leqslant 1$. Entonces, si $|(x, y)-(3, -1)| < \delta \leqslant 1$, $|x-3| < 1$ y

$$|x^2+2xy-3| < |x-3|+2|y+1|+4|x-3|+6|y+1| < 13\,\delta.$$

Por tanto, si escogemos $\delta = \text{mín}\,\{1, \varepsilon/13\}$

$$|x^2+2xy-3| < \varepsilon$$

siempre que $0 < |(x, y)-(3, -1)| < \delta$. Lo que prueba que 3 es el límite de f en $(3, -1)$.

Dimos el ejemplo 4.2 simplemente para ilustrar la definición de límite. Es fácil determinar este límite una vez que se ha dado un tratamiento sistemático de los límites. Es lo que vamos a hacer ahora. Comenzamos por determinar los límites de algunas funciones básicas.

4.3 Ejemplo. Si c es una función constante de R^n en R y $\mathbf{a}$ es un punto cualquiera en R^n, demuéstrese que $\lim\limits_{\mathbf{a}} c = c$.

SOLUCIÓN. Tómese $\varepsilon > 0$. Queremos demostrar que existe una $\delta > 0$ tal

que $|c-c| < \varepsilon$ siempre que $0 < |\mathbf{x}-\mathbf{a}| < \delta$. Es claro que podemos tomar como δ un número positivo cualquiera.

4.4 Ejemplo. Si I_k es una función proyección de R^n en R y $\mathbf{a}$ es un punto cualquiera en R^n, demuéstrese que $\lim\limits_{\mathbf{a}} I_k = a_k$.

SOLUCIÓN. Tómese $\varepsilon > 0$. Deseamos demostrar que hay una $\delta > 0$ tal que $|I_k(\mathbf{x})-a_k| = |x_k-a_k| < \varepsilon$ siempre que $0 < |\mathbf{x}-\mathbf{a}| < \delta$. Sea $\delta = \varepsilon$. Entonces, como $|x_k-a_k| \leqslant |\mathbf{x}-\mathbf{a}|$,

$$0 < |\mathbf{x}-\mathbf{a}| < \delta \quad \text{implica} \quad |x_k-a_k| < \varepsilon.$$

Damos ahora algunos teoremas sobre límites de combinaciones de funciones bajo las operaciones definidas en la sección anterior.

4.5 Teorema. *Si f y g son funciones de* R^n *en* R *tales que* $\lim\limits_{\mathbf{a}} f$ *y* $\lim\limits_{\mathbf{a}} g$ *existen y si* $\mathbf{a}$ *es un punto de acumulación de* $\mathscr{D}_f \cap \mathscr{D}_g$, *entonces*

$$\lim_{\mathbf{a}} (f+g) = \lim_{\mathbf{a}} f + \lim_{\mathbf{a}} g$$

$$\lim_{\mathbf{a}} (f-g) = \lim_{\mathbf{a}} f - \lim_{\mathbf{a}} g$$

$$\lim_{\mathbf{a}} (fg) = (\lim_{\mathbf{a}} f)(\lim_{\mathbf{a}} g)$$

$$\lim_{\mathbf{a}} (f/g) = (\lim_{\mathbf{a}} f)/(\lim_{\mathbf{a}} g), \quad (si \lim_{\mathbf{a}} g \neq 0).$$

Omitimos la prueba de este teorema, ya que es la misma que la del teorema correspondiente para funciones de R en R (problema 4).

4.6 Teorema. *Si f es una función de* R^n *en* R *tal que* $\lim\limits_{\mathbf{a}} f = b$, *$g$ es una función de* R *en* R *que es continua en b, y si* $\mathbf{a}$ *es un punto de acumulación del dominio de $g \circ f$, entonces*

$$\lim_{\mathbf{a}} (g \circ f) = g(b).$$

PRUEBA. Aunque la prueba de este teorema es la misma que la prueba del teorema correspondiente para funciones de R en R, la damos aquí de nuevo. Tomemos $\varepsilon > 0$. Como g es continua en b, existe un número $\eta > 0$ tal que

$$|g(y)-g(b)| < \varepsilon \text{ siempre que } y \in \mathscr{D}_g \text{ y } |y-b| < \eta.$$

Como $\lim\limits_{\mathbf{a}} f = b$, existe un número $\delta > 0$ tal que $|f(\mathbf{x})-b| < \eta$ siempre que $\mathbf{x} \in \mathscr{D}_f$ y $0 < |\mathbf{x}-\mathbf{a}| < \delta$.

Si $\mathbf{x} \in \mathscr{D}_{g \circ f}$ y $0 < |\mathbf{x} - \mathbf{a}| < \delta$, entonces $|f(\mathbf{x}) - b| < \eta$ y

$$|g(f(\mathbf{x})) - g(b)| < \varepsilon \quad \text{(figura 8)}.$$

Esto prueba que $\lim_{\mathbf{a}} (g \circ f) = g(b)$.

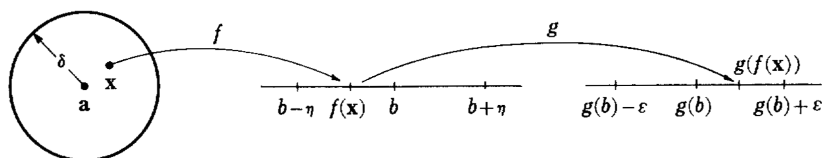


FIGURA 8

Podemos determinar los límites de las funciones polinomiales usando los ejemplos 4.3 y 4.4 y el teorema 4.5. Por ejemplo, si $f = I_1{}^2 + 2I_1 I_2$ (ejemplo 4.2), entonces

$$\begin{aligned}\lim_{(3,-1)} f &= \lim_{(3,-1)} I_1{}^2 + \lim_{(3,-1)} (2I_1 I_2) \\ &= (\lim_{(3,-1)} I_1)(\lim_{(3,-1)} I_1) + (\lim_{(3,-1)} 2)(\lim_{(3,-1)} I_1)(\lim_{(3,-1)} I_2) \\ &= (3)(3) + (2)(3)(-1) = 3.\end{aligned}$$

Usando los ejemplos 4.3 y 4.4 y el teorema 4.5 podemos, ciertamente, determinar el límite de cualquier función racional con tal de que el límite del denominador no sea cero. De acuerdo con el teorema 4.6, también a nuestra disposición, podemos manejar la mayoría de las funciones en que estamos interesados. Por ejemplo,

$$\lim_{(x,y)\to(2,3)} e^{xy} = e^6$$

ya que la función exponencial es continua en 6 y $\lim_{(x,y)\to(2,3)} xy = 6$.

Consideraremos ahora algunos ejemplos en que tenemos qué determinar el límite de una función racional, donde no podemos aplicar el teorema 4.5 por ser cero el límite del denominador.

Consideremos el límite en (0, 0) de la función f definida por $f(x, y) = \dfrac{1}{x^2 + y^2}$. Como $\lim_{(x,y)\to(0,0)} (x^2 + y^2) = 0$, no podemos aplicar el teorema 4.5. La curva de nivel de f correspondiente al valor c es el conjunto $\left\{(x, y) \,\middle|\, \dfrac{1}{x^2 + y^2} = c\right\}$. Si $c \leqslant 0$, este conjunto es vacío. Si $c > 0$, la curva de nivel es la circunferencia de ecuación $x^2 + y^2 = \dfrac{1}{c}$ (figura 9). Del diagrama

de las curvas de nivel se deduce que si (x, y) está próximo al origen, entonces $f(x, y)$ es grande. Parecería, pues, que $\lim\limits_{(0,0)} f$ no existe y, en realidad, $\lim\limits_{(0,0)} f = \infty$.

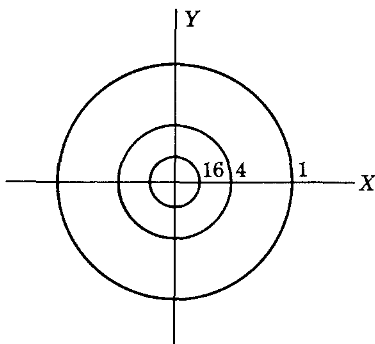


FIGURA 9

4.7 Definición. *La función f se dice que tiene **límite infinito en** $\mathbf{a}$, lo que se escribe* $\lim\limits_{\mathbf{a}} f = \infty$ *o* $\lim\limits_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = \infty$, *si* $\mathbf{a}$ *es un punto de acumulación de* $\mathscr{D}_f$ *y para cada número* $M > 0$ *hay un número* $\delta > 0$ *tal que*

$$f(\mathbf{x}) > M$$

siempre que $\mathbf{x} \in \mathscr{D}_f$ *y*

$$0 < |\mathbf{x} - \mathbf{a}| < \delta.$$

Nota. Si $\lim\limits_{\mathbf{a}} f = \infty$ seguiremos diciendo que no existe límite de f en $\mathbf{a}$ ya que ∞ no es un número real.

Si definimos una vecindad de infinito como un intervalo de la forma $\langle M, \infty\rangle$, entonces las definiciones 4.1 y 4.7 pueden considerarse casos especiales de la siguiente definición:

$\lim\limits_{\mathbf{a}} f = p$ *(donde p es un número real o ∞) si para toda vecindad $\mathscr{N}(p)$ de p existe una vecindad reducida $\mathscr{N}'(\mathbf{a})$ de $\mathbf{a}$ tal que*

$$f(\mathscr{N}'(\mathbf{a}) \cap \mathscr{D}_f) \subset \mathscr{N}(p).$$

Si una vecindad de $-\infty$ se define como un intervalo $\langle -\infty, M\rangle$, entonces, haciendo $p = -\infty$ en la anterior definición obtenemos una de $\lim\limits_{\mathbf{a}} f = -\infty$.

4.8 Ejemplo. Muéstrese que $\lim\limits_{(x,y)\to(0,0)} \dfrac{1}{x^2+y^2} = \infty$.

SOLUCIÓN. Tomemos un $M > 0$. Deseamos encontrar un número $\delta > 0$ tal que $\frac{1}{x^2+y^2} > M$ siempre que $0 < |\mathbf{x}| = \sqrt{x^2+y^2} < \delta$. Sea $\delta = \frac{1}{\sqrt{M}}$.

Consideraremos ahora el límite en $(0, 0)$ de la función f definida por $f(x, y) = \frac{x}{x^2+y^2}$. De nuevo no podemos aplicar el teorema 4.5 porque el límite del denominador es cero. En este caso el límite del numerador es también cero y no es claro cuál será el valor de f cerca del origen. La curva de nivel de f correspondiente a c es el conjunto $\left\{(x, y) \,\middle|\, \frac{x}{x^2+y^2} = c\right\}$. Si $c = 0$, la curva de nivel es el eje Y con el origen omitido. Si $c \neq 0$, la curva de nivel es la circunferencia $\left(x - \frac{1}{2c}\right)^2 + y^2 = \frac{1}{4c^2}$ con el origen omitido (figura 10). De la consideración de las curvas de nivel se deduce que f toma todos los valores reales en puntos arbitrariamente cercanos al origen. Parecería pues que $\lim\limits_{(0,0)} f$ no existe. Podemos demostrar que tal es el caso extendiendo la noción de límites derechos e izquierdos de las funciones de una variable real.

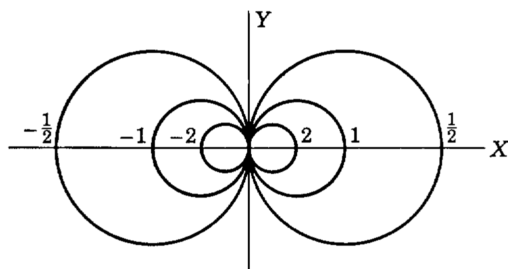


FIGURA 10

Si f es una función de R^n en R y $\mathscr{E}$ es un conjunto en R^n, sea $f_{\mathscr{E}}$ la función con dominio $\mathscr{E} \cap \mathscr{D}_f$ y regla de correspondencia

$$f_{\mathscr{E}}(\mathbf{x}) = f(\mathbf{x}) \quad \text{para} \quad \mathbf{x} \in \mathscr{E} \cap \mathscr{D}_f.$$

Entonces decimos que *el límite de la restricción de f a $\mathscr{E}$ en* $\mathbf{a}$ *es b*, lo que escribimos

$$\lim_{\mathbf{a}} f = b \text{ (sobre } \mathscr{E}) \quad \text{o} \quad \lim_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x}) = b \; (\mathbf{x} \in \mathscr{E} \cap \mathscr{D}_f),$$

si

$$\lim_{\mathbf{a}} f_{\mathscr{E}} = b.$$

De esta definición resulta claro que si $\lim_{\mathbf{a}} f = b$, entonces para cualquier conjunto $\mathscr{E}$ tal que $\mathbf{a}$ sea un punto de acumulación de $\mathscr{E} \cap \mathscr{D}_f$

$$\lim_{\mathbf{a}} f = b \quad (\text{sobre } \mathscr{E}).$$

Así pues, si hay alguna restricción de f que no tiene límite en $\mathbf{a}$ o hay dos restricciones de f que tienen en $\mathbf{a}$ límites distintos, ello indica que no existe el límite de f en $\mathbf{a}$.

4.9 Ejemplo. Si $f = \dfrac{I_1}{I_1{}^2 + I_2{}^2}$, pruébese que no existe $\lim_{\mathbf{0}} f$.

Solución. Si hacemos $\mathscr{E} = \{(x, y) \mid (x-h)^2 + y^2 = h^2\}$ —una curva de nivel de f— entonces

$$\lim_{(x,y)\to(0,0)} \frac{x}{x^2+y^2} = \lim_{(x,y)\to(0,0)} \frac{x}{2hx} = \frac{1}{2h} \quad (\text{sobre } \mathscr{E}).$$

Como los límites de f en el origen cuando lo restringimos a diferentes circunferencias que pasan por el origen son distintos, de acuerdo a lo que hemos visto, $\lim_{\mathbf{0}} f$ no existe.

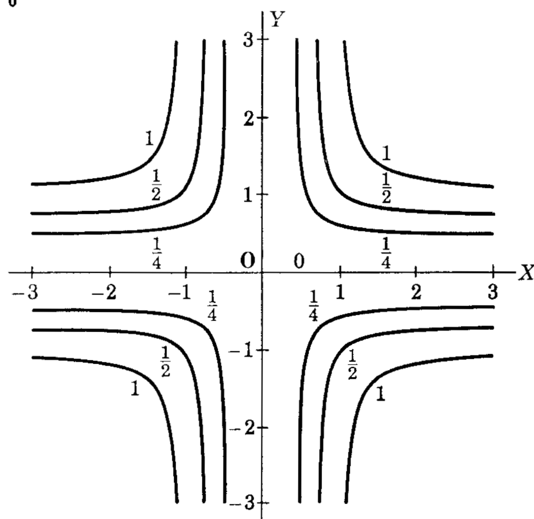


FIGURA 11

Como último ejemplo, consideremos el límite en el origen de la función $f = \dfrac{I_1{}^2 I_2{}^2}{I_1{}^2 + I_2{}^2}$. Tanto el numerador como el denominador tienen cero

como límite en el origen. Las curvas de nivel de f son los conjuntos $\left\{(x, y) \,\middle|\, \frac{x^2 y^2}{x^2+y^2} = c\right\}$. Si $c < 0$, el conjunto es vacío. Si $c = 0$ la curva de nivel consiste en los ejes X y Y sin el origen. Si $c > 0$, la curva de nivel tiene la ecuación $(x^2-c)(y^2-c) = c^2$ (figura 11). Del diagrama de curvas de nivel parece resultar que $\lim_{\mathbf{0}} f = 0$.

4.10 Ejemplo. Si $f = \frac{I_1{}^2 I_2{}^2}{I_1{}^2+I_2{}^2}$, demuéstrese que $\lim_{\mathbf{0}} f = 0$.

Solución. Tómese $\varepsilon > 0$. Si $x \neq 0$, entonces

$$\frac{x^2 y^2}{x^2+y^2} \leqslant \frac{x^2 y^2}{x^2} = y^2 < \varepsilon \quad \text{si} \quad 0 < x^2+y^2 < \varepsilon .$$

Si $x = 0$, entonces

$$\frac{x^2 y^2}{x^2+y^2} = 0 < \varepsilon \quad \text{si} \quad 0 < x^2+y^2 .$$

Así pues, si $\delta = \sqrt{\varepsilon}$, $0 < \sqrt{x^2+y^2} < \delta$ implica $\frac{x^2 y^2}{x^2+x^2} < \varepsilon$ y esto nos demuestra que $\lim_{\mathbf{0}} f = 0$.

Otro tipo de límite que algunas veces se presenta en la consideración de funciones reales de un vector es el de límite iterado, tal como, por ejemplo, $\lim_{x \to x_0} \lim_{y \to y_0} f(x, y)$. Este límite iterado tiene el significado siguiente: Para cada x fijo en una vecindad reducida $\mathscr{S}'(x_0; r)$ de x_0, tómese el límite de la función g en y_0 donde $g(y) = f(x, y)$. Si este límite existe para cada x en $\mathscr{S}'(x_0; r)$, entonces la función h, definida por $h(x) = \lim_{y \to y_0} f(x, y)$, existe en $\mathscr{S}'(x_0; r)$. El límite de h en x_0, si este límite existe, es $\lim_{x \to x_0} \lim_{y \to y_0} f(x, y)$.

4.11 Ejemplo. Encuéntrese $\lim_{x \to 3} \lim_{y \to 2} (x^2 y + 2xy^2)$.

Solución. $\lim_{x \to 3} \lim_{y \to 2} (x^2 y + 2xy^2) = \lim_{x \to 3} (2x^2 + 8x) = 42$.

Nótese que $\lim_{(x,y) \to (3,2)} (x^2 y + 2xy^2) = 42$ también. El siguiente teorema establece una relación entre $\lim_{x \to x_0} \lim_{y \to y_0} f(x, y)$ y $\lim_{(x,y) \to (x_0, y_0)} f(x, y)$.

4.12 Teorema. *Si* $\lim_{(x,y) \to (x_0, y_0)} f(x, y)$ *existe, y si, para cada* x *en una vecindad reducida de* x_0, $\lim_{y \to y_0} f(x, y)$ *existe, entonces*

$$\lim_{x \to x_0} \lim_{y \to y_0} f(x, y) = \lim_{(x,y) \to (x_0, y_0)} f(x, y) .$$

PRUEBA. Sea $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = b$ y $\lim_{y\to y_0} f(x,y) = h(x)$. La función h está definida para todo x en una vecindad reducida $\mathcal{N}'(x_0)$ de x_0. Tomemos $\varepsilon > 0$. Deseamos demostrar que existe una $\delta > 0$ tal que

$$|h(x)-b| < \varepsilon$$

siempre que $x \in \mathcal{N}'(x_0)$ y $|x-x_0| < \delta$. Como $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = b$, existe una $S > 0$ tal que

$$|f(x,y)-b| < \varepsilon/2$$

siempre que $(x,y) \in \mathcal{S}'((x_0,y_0);\delta) \cap \mathcal{D}_f$. Como $\lim_{y\to y_0} f(x,y) = h(x)$, para cualquier $x \in \mathcal{N}'(x_0)$ tal que $|x-x_0| < \delta$, existe un número y (figura 12) tal que $(x,y) \in \mathcal{S}'((x_0,y_0);\delta) \cap \mathcal{D}_f$ y

$$|f(x,y)-h(x)| < \varepsilon/2.$$

Por tanto, para $x \in \mathcal{N}'(x_0)$ y $|x-x_0| < \delta$ (y para una y adecuada)

$$|h(x)-b| \leqslant |h(x)-f(x,y)|+|f(x,y)-b| < \varepsilon/2+\varepsilon/2 = \varepsilon.$$

Así pues, $\lim_{x\to x_0} \lim_{y\to y_0} f(x,y) = b = \lim_{(x,y)\to(x_0,y_0)} f(x,y)$.

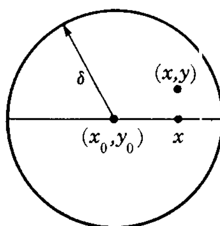


FIGURA 12

El teorema 4.12 implica que: si $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ existe y si, para cada x en una vecindad reducida de x_0, $\lim_{y\to y_0} f(x,y)$ existe y si, para cada y en una vecindad reducida de y_0, $\lim_{x\to x_0} f(x,y)$ existe, entonces

$$\lim_{x\to x_0} \lim_{y\to y_0} f(x,y) = \lim_{(x,y)\to(x_0,y_0)} f(x,y) = \lim_{y\to y_0} \lim_{x\to x_0} f(x,y).$$

Este resultado puede ser útil para demostrar que un límite no existe. Por ejemplo, si $f(x,y) = \dfrac{x}{x^2+y^2}$ (ejemplo 4.9), entonces para cada x en una vecindad reducida de 0, $\lim_{y\to 0} f(x,y) = \dfrac{1}{x}$. Entonces, si existe

$\lim_{(x,y)\to(0,0)} f(x, y)$, también existe $\lim_{x\to 0} \lim_{y\to 0} f(x, y)$. Sin embargo, $\lim_{x\to 0} \frac{1}{x}$ no existe y, por tanto, $\lim_{(x,y)\to(0,0)} f(x, y)$ no existe.

Problemas

1. Úsese la definición de límite para verificar que

a) $\lim_{(x,y)\to(1,2)} (x+y^2) = 5$

b) $\lim_{(x,y)\to(-3,2)} (xy+3y) = 0.$

2. Pruébese que si $\lim_{\mathbf{a}} f$ existe, entonces es único. Es decir, si $\lim_{\mathbf{a}} f = L_1$ y $\lim_{\mathbf{a}} f = L_2$, entonces $L_1 = L_2$.

3. Determínense los siguientes límites:

a) $\lim_{(x,y)\to(2,1)} xy$

b) $\lim_{(x,y,z)\to(3,-1,1)} x+y+z$

c) $\lim_{(2,1)} \operatorname{sen} \circ (I_1 I_2)$

d) $\lim_{(-2,3,1)} \dfrac{I_1{}^2 I_3{}^3 + I_2}{I_1{}^2 + I_2{}^2}$

e) $\lim_{(x,y)\to(3,5)} x^2+xy-y$

f) $\lim_{(3,2)} \ln \circ (I_1{}^2 I_2)$

g) $\lim_{(x,y)\to(0,0)} \dfrac{y \operatorname{sen} x}{x}$

h) $\lim_{(x,y)\to(0,0)} \dfrac{x \tan y}{y}$.

4. Si $\lim_{\mathbf{a}} f = L_1$ y $\lim_{\mathbf{a}} g = L_2$, pruébese que

a) $\lim_{\mathbf{a}} (f+g) = L_1 + L_2$

b) $\lim_{\mathbf{a}} (cf) = cL_1$

c) $\lim_{\mathbf{a}} (f-g) = L_1 - L_2$

d) $\lim_{\mathbf{a}} f^2 = L^2$ (*Sugerencia:* $f^2 = I^2 \circ f$, úsese el teorema 4.6.)

e) $\lim_{\mathbf{a}} fg = L_1 L_2$ (*Sugerencia:* $fg = \frac{1}{4}[(f+g)^2 - (f-g)^2]$.)

f) $\lim_{\mathbf{a}} \dfrac{1}{ag} = \dfrac{1}{L_2}$ $(L_2 \neq 0)$ $\left(\textit{Sugerencia:} \dfrac{1}{g} = I^{-1} \circ g, \text{ úsese el teorema 4.6.}\right)$

g) $\lim_{\mathbf{a}} \dfrac{f}{g} = \dfrac{L_1}{L_2}$ $(L_2 \neq 0)$.

5. Si $\lim_{\mathbf{a}} f = L$ demuéstrese que $\lim_{\mathbf{a}} f = L$. Demuéstrese con un ejemplo que lo recíproco no es cierto.

6. Demuéstrese que $\lim_{\mathbf{a}} |f| = 0$ implica que $\lim_{\mathbf{a}} f = 0$.

7. Determínese si cada una de las siguientes funciones tiene un límite en (0, 0).

a) $f(x, y) = \dfrac{x}{x+y}$ *b*) $f(x, y) = \dfrac{x^2}{x+y}$

c) $f(x, y) = \dfrac{x}{|x|+|y|}$ *d*) $f(x, y) = \dfrac{x^2}{|x|+|y|}$

e) $f(x, y) = \dfrac{x^2}{x^2+y^2}$ *f*) $f(x, y) = \dfrac{xy}{x^2+y^2}$

g) $f(x, y) = \dfrac{x^3}{x^2+y^2}$ *h*) $f(x, y) = \dfrac{y^2}{y^2+x^4}$.

8. Si existe una vecindad $\mathscr{S}(\mathbf{a}; r)$ de $\mathbf{a}$ tal que

$$f(\mathbf{x}) \leqslant g(\mathbf{x}) \leqslant h(\mathbf{x}) \quad \text{para toda} \quad \mathbf{x} \in \mathscr{S}'(\mathbf{a}; r)$$

y si

$$\lim_{\mathbf{a}} f = L = \lim_{\mathbf{a}} h,$$

demuéstrese que $\lim_{\mathbf{a}} g$ existe y $\lim_{\mathbf{a}} g = L$.

9. Si $\lim_{\mathbf{h}\to 0} f(\mathbf{h}) = L$ y $\mathbf{u}$ es un vector unitario fijo, demuéstrese que $\lim_{h\to 0} f(h) = L$.

10. Determínese $\lim_{x\to x_0} \lim_{y\to y_0} f(x, y)$, $\lim_{y\to y_0} \lim_{x\to x_0} f(x, y)$, y $\lim_{(x,y)\to(x_0, y_0)} f(x, y)$ cuando

a) $f(x, y) = x^3+xy^2$, $(x_0, y_0) = (-1, 3)$

b) $f(x, y) = x \operatorname{sen} xy$, $(x_0, y_0) = \left(1, \dfrac{\pi}{2}\right)$

c) $f(x, y) = \begin{cases} xy \operatorname{sen} \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$ $(x_0, y_0) = (0, 0)$.

11. Si $f(x, y) = \dfrac{x^2 - y^2}{x^2 + y^2}$, determínese $\lim_{x \to 0} \lim_{y \to 0} f(x, y)$ y $\lim_{y \to 0} \lim_{x \to 0} f(x, y)$. ¿Qué puede decirse sobre $\lim_{(x,y) \to (0,0)} f(x, y)$?

12. Si $f(x, y) = x \operatorname{sen} \dfrac{1}{y}$, determínense $\lim_{(x,y) \to (0,0)} f(x, y)$, $\lim_{x \to 0} \lim_{y \to 0} f(x, y)$, y $\lim_{y \to 0} \lim_{x \to 0} f(x, y)$ si existen.

5. CONTINUIDAD

5.1 Definición. *La función f es **continua en el punto** $\mathbf{a}$ de $\mathscr{D}_f$ si para cada $\varepsilon > 0$ existe una $\delta > 0$ tal que*

$$|f(\mathbf{x}) - f(\mathbf{a})| < \varepsilon$$

siempre que $\mathbf{x} \in \mathscr{D}_f$ y $|\mathbf{x} - \mathbf{a}| < \delta$.

En el lenguaje de las vecindades esta definición puede enunciarse como sigue: *f es continua en $\mathbf{a} \in \mathscr{D}_f$ si para cada vecindad $\mathscr{N}$ de $f(\mathbf{a})$ existe una vecindad $\mathscr{M}$ de $\mathbf{a}$ tal que*

$$f(\mathscr{M} \cap \mathscr{D}_f) \subset \mathscr{N}.$$

Si $\mathbf{a}$ pertenece a $\mathscr{D}_f$, pero no está en un punto de acumulación de $\mathscr{D}_f$, entonces f es continua en $\mathbf{a}$, pues en este caso existe una vecindad $\mathscr{M}$ de $\mathbf{a}$ tal que $\mathscr{M} \cap \mathscr{D}_f = \{\mathbf{a}\}$. Entonces, si $\mathscr{N}$ es una vecindad cualquiera de $f(\mathbf{a})$,

$$f(\mathscr{M} \cap \mathscr{D}_f) = \{f(\mathbf{a})\} \subset \mathscr{N}.$$

Si $\mathbf{a}$ es un punto de acumulación de $\mathscr{D}_f$, entonces la definición 5.1 es equivalente a la función f es continua en el punto $\mathbf{a}$ de $\mathscr{D}_f$ si

$$\lim_{\mathbf{a}} f = f(\mathbf{a}).$$

Correspondiéndose con el teorema sobre límites 4.5, tenemos el siguiente teorema sobre continuidad.

5.2 Teorema. *Si las funciones f y g son continuas en $\mathbf{a}$, entonces $f+g$, $f-g$, y fg son continuas en $\mathbf{a}$ y f/g es continua en $\mathbf{a}$ siempre que $g(\mathbf{a}) \neq 0$.*

PRUEBA. Si $\mathbf{a}$ no es un punto de acumulación de $\mathscr{D}_f \cap \mathscr{D}_g$, entonces estas funciones son todas continuas en $\mathbf{a}$. Si $\mathbf{a}$ es un punto de acumulación de

$\mathscr{D}_f \cap \mathscr{D}_g$, entonces **a** es un punto de acumulación tanto de $\mathscr{D}_f$ como de $\mathscr{D}_g$ y el teorema sigue del teorema 4.5. Por ejemplo,

$$\begin{aligned}\lim_{\mathbf{a}} (f+g) &= \lim_{\mathbf{a}} f + \lim_{\mathbf{a}} g \\ &= f(\mathbf{a})+g(\mathbf{a}) = [f+g](\mathbf{a}).\end{aligned}$$

Como $\lim_{\mathbf{a}} c = c$ y $\lim I_k = a_k = I_k(\mathbf{a})$, las funciones constantes y las funciones proyección son continuas en cualquier punto **a**. Así pues, según el teorema 5.2, las funciones polinomiales son continuas en todos los puntos de R^n; las funciones racionales son continuas en todos los puntos de R^n en que el denominador es distinto de cero, es decir, en todos los puntos en que las funciones están definidas.

5.3 Teorema. *Si f es una función de* R^n *en* R *que es continua en* **a** *y g es una función de* R *en* R *que es continua en $f(\mathbf{a})$, entonces $g \circ f$ es continua en* **a**.

PRUEBA. Si **a** no es un punto de acumulación de $\mathscr{D}_{g \circ f}$, entonces $g \circ f$ es continua en **a**. Si **a** es un punto de acumulación de $\mathscr{D}_{g \circ f}$, entonces, como $\mathscr{D}_{g \circ f} \subset \mathscr{D}_f$, **a** debe ser un punto de acumulación de $\mathscr{D}_f$ y $\lim_{\mathbf{a}} f = f(\mathbf{a})$. Luego de acuerdo con el teorema 4.6

$$\lim_{\mathbf{a}} (g \circ f) = g(f(\mathbf{a})) = [g \circ f](\mathbf{a}).$$

5.4 Ejemplo. Demuéstrese que la función f definida por $f(x, y) = \operatorname{sen} xy$ es continua en todos los puntos de R^2.

SOLUCIÓN. La función f es $\operatorname{sen} \circ (I_1 I_2)$. Como $I_1 I_2$ es una función polinomial, es continua en todos los puntos de R^2; la función seno es continua en todos los puntos de R. Luego, según el teorema 5.3, f es continua en todos los puntos de R^2.

La noción básica en continuidad es la de continuidad en un punto. Pero también usamos la terminología: "f es continua" o "f es continua sobre un conjunto $\mathscr{S}$". Definimos ahora estos términos.

5.5 Definición. *Una función es **continua** si es continua en cada punto de su dominio.*

5.6 Definición. *Una función f es **continua sobre un conjucto** $\mathscr{S} \subset \mathscr{D}_f$ si la función restringida $f_{\mathscr{S}}$ es continua.*

Así pues, podemos decir que la función f del ejemplo 5.4 es continua, o, equivalentemente, que es continua en R^2.

Toda función racional es continua. Obsérvese, sin embargo, que la función racional $f = \dfrac{1}{I_1^{\,2}+I_2^{\,2}}$ no es continua sobre el conjunto

$$\mathscr{S}(\mathbf{0}; 1) = \{(x, y) \mid x^2+y^2 < 1\},$$

ya que $\mathscr{S}(\mathbf{0}; 1)$ no está contenida en $\mathscr{D}_f$.

Una función real de una variable real continua sobre un intervalo posee la propiedad del valor intermedio: si f es continua sobre $[a, b]$ y $f(a) < f(b)$ y t es un número cualquiera tal que $f(a) < t < f(b)$, entonces existe un punto $c \in \langle a, b \rangle$ tal que $f(c) = t$. Demostraremos que las funciones reales de un vector poseen también esta propiedad. Introduciremos primero algunos términos de la terminología usual.

Si $\mathscr{E} \subset \mathscr{F} \subset \mathrm{R}^n$, entonces decimos que $\mathscr{E}$ es *abierto respecto a* $\mathscr{F}$ si existe un conjunto abierto $\mathscr{G}$ en R^n tal que $\mathscr{E} = \mathscr{G} \cap \mathscr{F}$. Así pues, $\mathscr{E}$ es abierto respecto a $\mathscr{F}$ si y sólo si para todo $\mathbf{x} \in \mathscr{E}$ hay una vecindad $\mathscr{S}(\mathbf{x}; r)$ de $\mathbf{x}$ tal que $\mathscr{S}(\mathbf{x}; r) \cap \mathscr{F} \subset \mathscr{E}$. Por ejemplo, en R el conjunto $[0, 1\rangle$ es abierto relativo a $[0, \infty\rangle$ ya que $[0, 1\rangle = \langle -1, 1\rangle \cap [0, \infty\rangle$.

Un conjunto $\mathscr{E} \subset \mathrm{R}^n$ se dice que es *conexo* si no existen dos conjuntos no vacíos ajenos $\mathscr{A}$ y $\mathscr{B}$ ambos abiertos respecto a $\mathscr{E}$ tales que $\mathscr{E} = \mathscr{A} \cup \mathscr{B}$. Si $\mathscr{E}$ es abierto, entonces $\mathscr{E}$ es conexo si no puede representarse como la unión de dos conjuntos ajenos, no vacíos y abiertos. Por ejemplo, un intervalo es un conjunto conexo en R, un círculo y un rectángulo son conjuntos conexos en R^2, y una vecindad es un conjunto conexo en R^n. Un conjunto de la forma $\{\mathbf{x} \mid |\mathbf{x}| > 1\}$ no es conexo en R, pero es conexo si es un conjunto en R^n para $n > 1$.

5.7 Teorema. (Teorema del valor intermedio.) *Sea f una función real continua sobre un conjunto abierto y conexo $\mathscr{E} \subset \mathrm{R}^n$ y sea $f(\mathbf{a}) < f(\mathbf{b})$ para algunos $\mathbf{a}, \mathbf{b} \in \mathscr{E}$. Para cada t tal que $f(\mathbf{a}) < t < f(\mathbf{b})$ hay un punto $\mathbf{c} \in \mathscr{E}$ tal que $f(\mathbf{c}) = t$.*

PRUEBA. Sea $\mathscr{A} = \{\mathbf{x} \mid \mathbf{x} \in \mathscr{E} \text{ y } f(\mathbf{x}) < t\}$ y $\mathscr{B} = \{\mathbf{x} \mid \mathbf{x} \in \mathscr{E} \text{ y } f(\mathbf{x}) > t\}$. Claramente $\mathscr{A}$ y $\mathscr{B}$ son ajenos y no vacíos: $\mathbf{a} \in \mathscr{A}$ y $\mathbf{b} \in \mathscr{B}$. Sea $\mathbf{x} \in \mathscr{A}$. Como $f(\mathbf{x}) < t$ y f es continua en $\mathbf{x}$, existe una vecindad $\mathscr{S}(\mathbf{x}) \subset \mathscr{E}$ tal que $f(\mathbf{y}) < t$ para todo $\mathbf{y} \in \mathscr{S}(\mathbf{x})$. Luego $\mathscr{A}$ es abierto. Análogamente puede verse que $\mathscr{B}$ es también abierto. Pero $\mathscr{E}$ es conexo. Luego no puede ser la unión de dos conjuntos ajenos, no vacíos y abiertos $\mathscr{A}$ y $\mathscr{B}$. Luego hay al menos un punto $\mathbf{c} \in \mathscr{E}$ tal que $\mathbf{c} \notin \mathscr{A} \cup \mathscr{B}$, es decir, tal que $f(\mathbf{c}) = t$.

Problemas

1. Si f es continua en $\mathbf{a}$ y $b < f(\mathbf{a}) < c$, demuéstrese que existe una vecindad $\mathscr{S}(\mathbf{a}; \delta)$ de $\mathbf{a}$ tal que $b < f(\mathbf{x}) < c$ para todo $\mathbf{x} \in \mathscr{D}_f \cap \mathscr{S}(\mathbf{a}; \delta)$.

2. Úsese la definición 5.1 para verificar que la función $f = I_1^{\,3}+I_2^{\,3}$ es continua en $(0, 0)$ y en $(2, 1)$.

3. Determínese el conjunto de puntos en que f es continua cuando

a) $f(x, y, z) = \dfrac{2xyz}{x-y}$ *b*) $f(x, y) = \tan xy$

c) $f = \dfrac{1}{I_1{}^2 + 2I_1 + I_2{}^2}$ *d*) $f = \exp \circ (I_1 + I_2)$.

4. ¿Es continua en (0, 0) la función f definida por

$$\begin{cases} f(x, y) = \dfrac{y^3}{x^2+y^2}, & (x, y) \neq (0, 0)? \\ f(0, 0) = 0 \end{cases}$$

5. ¿Es continua en (0, 0) la función f definida por

$$\begin{cases} f(x, y) = \dfrac{y^2}{x^2+y^2}, & (x, y) \neq (0, 0)? \\ f(0, 0) = 0 \end{cases}$$

6. Supongamos que f es una función de R^3 en R y que f es continua en el punto (x_0, y_0, z_0). Definamos la función g de R^2 en R de acuerdo con la regla

$$g(x, y) = f(x, y, z_0).$$

Demuéstrese que g es continua en el punto (x_0, y_0).

7. Si **a** es un punto de acumulación de un conjunto $\mathscr{E}$, demuéstrese que $\mathbf{a} \in \overline{\mathscr{E}}$.

8. Demuéstrese que si $\mathscr{E}$ es abierta entonces todo punto de $\mathscr{E}$ es un punto de acumulación de $\mathscr{E}$.

***9.** Demuéstrese que un conjunto $\mathscr{E}$ en R es conexo si y sólo si $\mathscr{E}$ es un intervalo.[1]

6. FUNCIONES DIFERENCIABLES

Supongamos que f es una función real de variable real definida sobre un intervalo abierto $\mathscr{I}$. Si la derivada de f existe en el punto $x \in \mathscr{I}$, entonces

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

[1] Volumen 1, pág. 434, teorema 4.5

Haciendo $\varphi(x;h) = \frac{1}{h}[f(x+h) - f(x)] - f'(x)$ cuando $h \neq 0$, obtenemos

6.1 $$f(x+h) = f(x)+f'(x)h+\varphi(x;h)h \quad \text{donde} \quad \lim_{h\to 0} \varphi(x;h) = 0.$$

La relación 6.1 se verifica para todos los valores distintos de cero de h tales que $x+h\in\mathscr{J}$.

Decimos que la función f de R en R es diferenciable en el punto x si f está definida en una vecindad $\mathscr{S}(x;r)$ de x y si existe un número a (independiente de h) tal que para cualquier punto $x+h$ en $\mathscr{S}'(x;r)$

$$f(x+h) = f(x)+ah+\varphi(x;h)h \quad \text{donde} \quad \lim_{h\to 0} \varphi(x;h) = 0.$$

El término ah se llama diferencial de f en x y h y se representa por $df(x;h)$. Hemos demostrado en los anteriores renglones que si f tiene una derivada en el punto x, entonces f es diferenciable en x con $a = f'(x)$ y, por tanto, $df(x;h) = f'(x)h$.

Probaremos ahora que si la función f de R en R es diferenciable en x, entonces f tiene una derivada en x. Como para todo $x+h$ en una vecindad reducida de x,

$$f(x+h) = f(x)+ah+\varphi(x;h)h \quad \text{donde} \quad \lim_{h\to 0} \varphi(x;h) = 0,$$

tenemos

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = a.$$

Por tanto, $f'(x)$ existe y es igual a a.

Análogamente, una función de R en R^n es diferenciable en un punto si y sólo si tiene una derivada en el punto. En el caso de funciones de R^n en R no existe una extensión directa de la definición común de derivada. Sin embargo, como la notación de diferenciabilidad se extiende fácilmente, tomamos esta noción como básica y definimos la derivada en este contexto.

6.2 Definición. *Si f es una función de* R^n *en* R, *decimos que f es* ***diferenciable en el punto*** $\mathbf{x}$ *si f está definida en una vecindad* $\mathscr{S}(\mathbf{x};r)$ de $\mathbf{x}$ *y si existe un vector* $\mathbf{a}$ *(independiente de* $\mathbf{h}$*) tal que para cualquier punto* $\mathbf{x}+\mathbf{h}$ en $\mathscr{S}'(\mathbf{x};r)$

6.3 $$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x})+\mathbf{a}\cdot\mathbf{h}+\boldsymbol{\varphi}(\mathbf{x};\mathbf{h})\cdot\mathbf{h} \quad \textit{donde} \quad \lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\varphi}(\mathbf{x};\mathbf{h}) = \mathbf{0}.$$

El término $\mathbf{a}\cdot\mathbf{h}$ *se llama* ***diferencial*** *de f en* $\mathbf{x}$ *y* $\mathbf{h}$ *y se denota por* $df(\mathbf{x};\mathbf{h})$. *El vector* $\mathbf{a}$ *se llama* ***derivada*** *de f en* $\mathbf{x}$ *y se denota por* $\mathbf{D}f(\mathbf{x})$.

Nótese que $\boldsymbol{\varphi}(\mathbf{x};\mathbf{h})$, para un $\mathbf{x}$ fijo, es una función de R^n en R^n. Consideraremos tales funciones en el próximo capítulo, pero ahora debemos

explicar qué es lo que queremos decir por $\lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) = \mathbf{0}$. Si escribimos

$$\boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) = (\varphi_1(\mathbf{x}; \mathbf{h}), \ldots, \varphi_n(\mathbf{x}; \mathbf{h})),$$

entonces $\lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\varphi}(\mathbf{x}, \mathbf{h}) = \mathbf{0}$ quiere decir $\lim_{\mathbf{h}\to\mathbf{0}} \varphi_k(\mathbf{x}; \mathbf{h}) = 0$ para $k = 1, \ldots, n$.

6.4 Ejemplo. Demuéstrese que la función f definida por

$$f(x, y) = x^2 + xy$$

es diferenciable en todo punto de R^2.

SOLUCIÓN. Sea $\mathbf{x} = (x, y)$ un punto cualquiera y $\mathbf{h} = (h_1, h_2)$ un vector cualquiera de R^2. Entonces

$$\begin{aligned} f(\mathbf{x}+\mathbf{h}) &= f(x+h_1, y+h_2) = (x+h_1)^2 + (x+h_1)(y+h_2) \\ &= x^2 + 2xh_1 + h_1{}^2 + xy + yh_1 + xh_2 + h_1 h_2 \\ &= x^2 + xy + (2x+y, x)\cdot(h_1, h_2) + (h_1, h_1)\cdot(h_1, h_2) \\ &= f(\mathbf{x}) + \mathbf{a}\cdot\mathbf{h} + \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})\cdot\mathbf{h}, \end{aligned}$$

donde $\mathbf{a} = (2x+y, x)$ y $\boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) = (h_1, h_1)$.

Queda por probar que $\lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) = \mathbf{0}$. Pero es claro que

$$\lim_{\mathbf{h}\to\mathbf{0}} (h_1, h_1) = (\lim_{\mathbf{h}\to\mathbf{0}} h_1, \lim_{\mathbf{h}\to\mathbf{0}} h_1) = (0, 0).$$

Por tanto, f es diferenciable en todo punto de R^2. Además, la diferencial de f en $\mathbf{x}$ y $\mathbf{h}$ es

$$df(\mathbf{x}; \mathbf{h}) = (2x+y, x)\cdot(h_1, h_2) = (2x+y)h_1 + xh_2$$

y la derivada de f en $\mathbf{x}$ es

$$\mathbf{D}f(\mathbf{x}) = (2x+y, x).$$

Cuando decimos de una función que es diferenciable sobre un conjunto queremos decir que la función ha de estar definida en una vecindad de cada punto del conjunto. De aquí que consideremos solamente la diferenciabilidad sobre conjuntos abiertos.

6.5 Teorema. *Si f es diferenciable en $\mathbf{x}$, entonces es continua en $\mathbf{x}$.*

PRUEBA. Como f es diferenciable en $\mathbf{x}$, f está definida en una vecindad $\mathscr{S}(\mathbf{x}; r)$ de $\mathbf{x}$. Deseamos ahora probar que

$$\lim_{\mathbf{h}\to\mathbf{0}} f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}).$$

Para cualquier punto $\mathbf{x}+\mathbf{h}$ de $\mathscr{S}'(\mathbf{x}; r)$ tenemos

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x})+\mathbf{a}\cdot\mathbf{h}+\boldsymbol{\varphi}(\mathbf{x};\mathbf{h})\cdot\mathbf{h} \quad \text{donde} \quad \lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\varphi}(\mathbf{x};\mathbf{h}) = \mathbf{0}.$$

Por tanto,

$$\lim_{\mathbf{h}\to\mathbf{0}} f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x})$$

y f es continua en $\mathbf{x}$.

El recíproco de este teorema no se verifica. Por ejemplo, la función definida por $f(x, y) = \sqrt{x^2+y^2}$ es continua en el origen, pero no es diferenciable en el (problema 2).

Como en el caso de funciones reales de variable real, la suma, el producto y la composición de funciones diferenciables son diferenciables.

6.6 Teorema. *Si f y g son diferenciables en un punto $\mathbf{x}$ en R^n, entonces $f+g$ y fg son diferenciables en $\mathbf{x}$ y*

$$\begin{aligned} d[f+g](\mathbf{x};\mathbf{h}) &= df(\mathbf{x};\mathbf{h})+dg(\mathbf{x};\mathbf{h}) \\ \mathbf{D}[f+g](\mathbf{x}) &= \mathbf{D}f(\mathbf{x})+\mathbf{D}g(\mathbf{x}) \\ d[fg](\mathbf{x};\mathbf{h}) &= f(\mathbf{x})\,dg(\mathbf{x};\mathbf{h})+g(\mathbf{x})\,df(\mathbf{x};\mathbf{h}) \\ \mathbf{D}[fg](\mathbf{x}) &= f(\mathbf{x})\mathbf{D}g(\mathbf{x})+g(\mathbf{x})\mathbf{D}f(\mathbf{x}). \end{aligned}$$

PRUEBA. Probaremos el teorema solamente para el producto. La prueba para la suma se deja para el estudiante (problema 3). Como f y g son diferenciables en $\mathbf{x}$, existen vectores $\mathbf{a}$ y $\mathbf{b}$ y una vecindad $\mathscr{S}(\mathbf{x}; r)$ de $\mathbf{x}$ tales que para cualquier punto $\mathbf{x}+\mathbf{h}$ en $\mathscr{S}'(\mathbf{x}; r)$

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x})+\mathbf{a}\cdot\mathbf{h}+\boldsymbol{\varphi}(\mathbf{x};\mathbf{h})\cdot\mathbf{h} \quad \text{donde} \quad \lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\varphi}(\mathbf{x};\mathbf{h}) = \mathbf{0}.$$

y

$$g(\mathbf{x}+\mathbf{h}) = g(\mathbf{x})+\mathbf{b}\cdot\mathbf{h}+\boldsymbol{\psi}(\mathbf{x};\mathbf{h})\cdot\mathbf{h} \quad \text{donde} \quad \lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\psi}(\mathbf{x};\mathbf{h}) = \mathbf{0}.$$

Entonces, si $\mathbf{x}+\mathbf{h}\in\mathscr{S}'(\mathbf{x}; r)$

$$\begin{aligned} [fg](\mathbf{x}+\mathbf{h}) &= [f(\mathbf{x})+\mathbf{a}\cdot\mathbf{h}+\boldsymbol{\varphi}(\mathbf{x};\mathbf{h})\cdot\mathbf{h}]\,[g(\mathbf{x})+\mathbf{b}\cdot\mathbf{h}+\boldsymbol{\psi}(\mathbf{x};\mathbf{h})\cdot\mathbf{h}] \\ &= [fg](\mathbf{x})+[f(\mathbf{x})\mathbf{b}+g(\mathbf{x})\mathbf{a}]\cdot\mathbf{h}+\boldsymbol{\theta}(\mathbf{x};\mathbf{h})\cdot\mathbf{h} \end{aligned}$$

donde

$$\begin{aligned} \boldsymbol{\theta}(\mathbf{x};\mathbf{h}) = {} & f(\mathbf{x})\,\boldsymbol{\psi}(\mathbf{x};\mathbf{h})+(\mathbf{a}\cdot\mathbf{h})\,\mathbf{b}+(\mathbf{a}\cdot\mathbf{h})\,\boldsymbol{\psi}(\mathbf{x};\mathbf{h})+g(\mathbf{x})\,\boldsymbol{\varphi}(\mathbf{x};\mathbf{h}) \\ & +(\mathbf{b}\cdot\mathbf{h})\,\boldsymbol{\varphi}(\mathbf{x};\mathbf{h})+[\boldsymbol{\varphi}(\mathbf{x};\mathbf{h})\cdot\mathbf{h}]\,\boldsymbol{\psi}(\mathbf{x};\mathbf{h}). \end{aligned}$$

Como $\lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\theta}(\mathbf{x};\mathbf{h}) = \mathbf{0}$, fg es diferenciable en $\mathbf{x}$ y

$$\begin{aligned} d[fg](\mathbf{x};\mathbf{h}) &= f(\mathbf{x})\,dg(\mathbf{x};\mathbf{h})+g(\mathbf{x})\,df(\mathbf{x};\mathbf{h}) \\ \mathbf{D}[fg](\mathbf{x}) &= f(\mathbf{x})\mathbf{D}g(\mathbf{x})+g(\mathbf{x})\mathbf{D}f(\mathbf{x}). \end{aligned}$$

6.7 Teorema. *Si f es una función de R_n en R que es diferenciable en $\mathbf{x}$ y g es una función de R en R que es diferenciable en $f(\mathbf{x})$, entonces $g \circ f$ es diferenciable en $\mathbf{x}$ y*

$$d[g \circ f](\mathbf{x}; \mathbf{h}) = dg(f(\mathbf{x}); df(\mathbf{x}; \mathbf{h}))$$
$$\mathbf{D}[g \circ f](\mathbf{x}) = Dg(f(\mathbf{x}))\mathbf{D}f(\mathbf{x}).$$

PRUEBA. Como f es diferenciable en $\mathbf{x}$, existe una vecindad $\mathscr{S}(\mathbf{x}; r)$ de $\mathbf{x}$ y un vector $\mathbf{a}$ tales que para cualquier punto $\mathbf{x}+\mathbf{h}$ en $\mathscr{S}'(\mathbf{x}; r)$

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \mathbf{a}\cdot\mathbf{h} + \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})\cdot\mathbf{h} \quad \text{donde} \lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) = \mathbf{0}.$$

La diferenciabilidad de g en $f(\mathbf{x})$ implica que existe una vecindad $\mathscr{S}(f(\mathbf{x}); s)$ de $f(\mathbf{x})$ tal que para cualquier punto $f(\mathbf{x})+k$ en $\mathscr{S}'(f(\mathbf{x}); s)$

$$g(f(\mathbf{x})+k) = g(f(\mathbf{x})) + g'(f(\mathbf{x}))k + \psi(f(\mathbf{x}); k)k$$

donde $\lim_{k\to 0} \psi(f(\mathbf{x}); k) = 0$. Si hacemos $\psi(f(\mathbf{x}); 0) = 0$, entonces la expresión anterior se verifica para todos los puntos $f(\mathbf{x})+k$ en $\mathscr{S}(f(\mathbf{x}); s)$. Como f es continua en $\mathbf{x}$ podemos suponer que $\mathscr{S}(\mathbf{x}; r)$ está escogida de tal modo que $f(\mathscr{S}(\mathbf{x}; r)) \subset \mathscr{S}(f(\mathbf{x}); s)$. Luego si $\mathbf{x}+\mathbf{h} \in \mathscr{S}'(\mathbf{x}; r)$,

$$\begin{aligned}
[g \circ f](\mathbf{x}+\mathbf{h}) &= g(f(\mathbf{x}) + \mathbf{a}\cdot\mathbf{h} + \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})\cdot\mathbf{h}) \\
&= g(f(\mathbf{x})) + g'(f(\mathbf{x}))\,[\mathbf{a}\cdot\mathbf{h} + \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})\cdot\mathbf{h}] \\
&\quad + [\mathbf{a}\cdot\mathbf{h} + \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})\cdot\mathbf{h}]\,\psi(f(\mathbf{x}); \mathbf{a}\cdot\mathbf{h} + \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})\cdot\mathbf{h}) \\
&= g(f(\mathbf{x})) + g'(f(\mathbf{x}))\mathbf{a}\cdot\mathbf{h} + \boldsymbol{\theta}(\mathbf{x}; \mathbf{h})\cdot\mathbf{h}
\end{aligned}$$

donde

$$\boldsymbol{\theta}(\mathbf{x}; \mathbf{h}) = g'(f(\mathbf{x}))\,\boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) + [\mathbf{a} + \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})]\,\psi(f(\mathbf{x}); \mathbf{a}\cdot\mathbf{h} + \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})\cdot\mathbf{h}).$$

Como $\lim_{\mathbf{h}\to\mathbf{0}} \psi(f(\mathbf{x}); \mathbf{a}\cdot\mathbf{h} + \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})\cdot\mathbf{h}) = 0$ (teorema 4.6, pág. 175), $\lim_{\mathbf{h}\to 0} \boldsymbol{\theta}(\mathbf{x}; \mathbf{h}) = \mathbf{0}$ y, por tanto, $g \circ f$ es diferenciable en $\mathbf{x}$ y

$$d[g \circ f](\mathbf{x}; \mathbf{h}) = dg(f(\mathbf{x}); df(\mathbf{x}; \mathbf{h}))$$
$$\mathbf{D}[g \circ f](\mathbf{x}) = Dg(f(\mathbf{x}))\mathbf{D}f(\mathbf{x}).$$

6.8 Corolario. *Si f y g son diferenciables en un punto $\mathbf{x}$ de R^n y $g(\mathbf{x}) \neq 0$, entonces $\frac{f}{g}$ es diferenciable en $\mathbf{x}$ y*

$$d\left[\frac{f}{g}\right](\mathbf{x}; \mathbf{h}) = \frac{g(\mathbf{x})\,df(\mathbf{x}; \mathbf{h}) - f(\mathbf{x})\,dg(\mathbf{x}; \mathbf{h})}{g^2(\mathbf{x})}$$

$$\mathbf{D}\left[\frac{f}{g}\right](\mathbf{x}) = \frac{g(\mathbf{x})\,\mathbf{D}f(\mathbf{x}) - f(\mathbf{x})\,\mathbf{D}g(\mathbf{x})}{g^2(\mathbf{x})}.$$

PRUEBA Como $\frac{f}{g} = f \cdot \frac{1}{g}$,

$$d\left[\frac{f}{g}\right](\mathbf{x}; \mathbf{h}) = f(\mathbf{x})\, d\left[\frac{1}{g}\right](\mathbf{x}; \mathbf{h}) + \frac{1}{g(\mathbf{x})}\, df(\mathbf{x}; \mathbf{h}).$$

Además, $\frac{1}{g} = I^{-1} \circ g$ y, por tanto,

$$d\left[\frac{1}{g}\right](\mathbf{x}; \mathbf{h}) = dI^{-1}(g(\mathbf{x}); dg(\mathbf{x}; \mathbf{h})) = -I^{-2}(g(\mathbf{x}))\, dg(\mathbf{x}; \mathbf{h})$$

$$= -\frac{dg(\mathbf{x}; \mathbf{h})}{g^2(\mathbf{x})}.$$

Por tanto

$$d\left[\frac{f}{g}\right](\mathbf{x}; \mathbf{h}) = \frac{g(\mathbf{x})\, df(\mathbf{x}; \mathbf{h}) - f(\mathbf{x})\, dg(\mathbf{x}; \mathbf{h})}{g^2(\mathbf{x})}.$$

Como

$$d\left[\frac{f}{g}\right](\mathbf{x}; \mathbf{h}) = \frac{g(\mathbf{x})\, \mathbf{D}f(\mathbf{x}) \cdot \mathbf{h} - f(\mathbf{x})\, \mathbf{D}g(\mathbf{x}) \cdot \mathbf{h}}{g^2(\mathbf{x})}$$

$$= \frac{g(\mathbf{x})\, \mathbf{D}f(\mathbf{x}) - f(\mathbf{x})\, \mathbf{D}g(\mathbf{x})}{g^2(\mathbf{x})} \cdot \mathbf{h},$$

$$\mathbf{D}\left[\frac{f}{g}\right](\mathbf{x}) = \frac{g(\mathbf{x})\, \mathbf{D}f(\mathbf{x}) - f(\mathbf{x})\, \mathbf{D}g(\mathbf{x})}{g^2(\mathbf{x})}.$$

La prueba del siguiente teorema es análoga a la del teorema 6.7 y se deja para el estudiante (problema 4).

6.9 Teorema. *Si* $\mathbf{g}$ *es una función de* R *en* R^n *que es diferenciable en el punto* t *y* f *es una función de* R^n *en* R *que es diferenciable en* $\mathbf{g}(t)$, *entonces* $f \circ \mathbf{g}$ *es diferenciable en* t *y*

$$d[f \circ \mathbf{g}]\,(t; h) = df(\mathbf{g}(t); d\mathbf{g}(t; h))$$
$$D[f \circ \mathbf{g}]\,(t) = \mathbf{D}f(\mathbf{g}(t)) \cdot D\mathbf{g}(t).$$

Esta fórmula que acabamos de dar para la derivada de la composición de funciones se llama a veces *regla de la cadena.*

Usando el teorema 6.9 y el teorema del valor medio para funciones reales de variable real, podemos fácilmente derivar un teorema del valor medio para funciones reales de un vector.

6.10 Teorema. (Teorema del valor medio.) *Si f es diferenciable sobre un conjunto abierto $\mathscr{E}$ que contiene el segmento rectilíneo cerrado que va de* **x** *a* **y**, *entonces existe un número $\theta \in \langle 0, 1 \rangle$ tal que*

$$f(\mathbf{y}) - f(\mathbf{x}) = (\mathbf{y} - \mathbf{x}) \cdot \mathbf{D}f(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})).$$

PRUEBA. Sea $\mathbf{g}(t) = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$ y $F = f \circ \mathbf{g}$. Entonces

$$f(\mathbf{y}) - f(\mathbf{x}) = f(\mathbf{g}(1)) - f(\mathbf{g}(0)) = F(1) - F(0).$$

De acuerdo con el teorema 6.9, F es diferenciable sobre $[0, 1]$. Aplicando el teorema del valor medio para funciones reales de una variable real a F, obtenemos

$$F(1) - F(0) = F'(\theta) \quad \text{para algun} \quad \theta \in \langle 0, 1 \rangle.$$

Esto implica

$$\begin{aligned} f(\mathbf{y}) - f(\mathbf{x}) &= D[f \circ \mathbf{g}](\theta) = \mathbf{D}f(\mathbf{g}(\theta)) \cdot D\mathbf{g}(\theta) \\ &= (\mathbf{y} - \mathbf{x}) \cdot \mathbf{D}f(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})). \end{aligned}$$

Problemas

1. Demuéstrese que las siguientes funciones son diferenciables en cualquier punto **x** de R^2 y proporciónese el valor de la diferencial en **x** y **h** y el valor de la derivada en **x**.

a) $f(x, y) = x + y$ *b*) $f(x, y) = 2xy$
c) $f(x, y) = x^2 + y^2$ *d*) $f(x, y) = x^2 y + x^3$.

2. Demuéstrese que la función definida por $f(x, y) = \sqrt{x^2 + y^2}$ es continua en el origen, pero no es diferenciable allí.

3. Si f y g son diferenciables en **x**, demuéstrese que $f + g$ es diferenciable en **x** y que

$$\begin{aligned} d[f+g](\mathbf{x}; \mathbf{h}) &= df(\mathbf{x}; \mathbf{h}) + dg(\mathbf{x}; \mathbf{h}) \\ \mathbf{D}[f+g](\mathbf{x}) &= \mathbf{D}f(\mathbf{x}) + \mathbf{D}g(\mathbf{x}). \end{aligned}$$

4. Pruébese el teorema 6.9.

5. Demuéstrese que $\mathbf{D}c = \mathbf{0}$, donde c es una función constante de $\mathbf{R}^n$ en R.

6. Si I_1 e I_2 son las funciones proyección de R^2, demuéstrese que

a) $\mathbf{D}I_1 = (1, 0)$ *b*) $\mathbf{D}I_2 = (0, 1)$.

7. Determínese $\mathbf{D}f(x, y)$ si

a) $f(x, y) = 3x + 2y$ *b*) $f(x, y) = xy$
c) $f(x, y) = x^2$ *d*) $f(x, y) = x^2 y$
e) $f(x, y) = \dfrac{x^2}{y}$ *f*) $f(x, y) = \operatorname{sen}(xy)$.

8. Un conjunto $\mathscr{E}$ se dice que es *convexo* si, para dos puntos cualesquiera $\mathbf{x}$ y $\mathbf{y}$ de $\mathscr{E}$, el segmento rectilíneo de $\mathbf{x}$ a $\mathbf{y}$ se encuentra en $\mathscr{E}$.

a) Demuéstrese que una vecindad $\mathscr{S}(\mathbf{a}; r)$ es convexa.

b) Demuéstrese que si $\mathbf{D}f(\mathbf{x}) = \mathbf{0}$ en todos los puntos $\mathbf{x}$ de un conjunto abierto y convexo $\mathscr{E}$, entonces es una constante sobre $\mathscr{E}$.

9. *a*) Demuéstrese que si $\mathbf{D}f(\mathbf{x}) = \mathbf{0}$ en todo punto $\mathbf{x}$ de un conjunto abierto y conexo $\mathscr{E}$, entonces f es una constante sobre $\mathscr{E}$.

Sugerencia. Úsese el problema 8 para demostrar que $f^*(c) = \{\mathbf{x} \mid f(\mathbf{x}) = c\}$ es abierto.

b) Proporciónese un ejemplo de una función f definida en un conjunto abierto $\mathscr{E}$ de R^2 tal que $\mathbf{D}f(x, y) = \mathbf{0}$ para todo $(x, y) \in \mathscr{E}$ tal que f no sea constante sobre $\mathscr{E}$.

7. DERIVADAS DIRECCIONALES

A fin de disponer de una imagen geométrica sencilla, inicialmente restringiremos nuestra discusión a funciones de R^2 en R. Entonces la gráfica de f es una superficie en R^3 (figura 13a). En un punto $\mathbf{x}$ del dominio de f la función no tendrá en general una razón de cambio única sino que cambiará en proporciones distintas, según cuál sea la dirección en que $\mathbf{x}$ se mueva. La razón de cambio de f en la dirección $\mathbf{u}$, donde $\mathbf{u}$ es un vector unitario, está dada por

7.1 $$\lim_{h \to 0} \frac{f(\mathbf{x}+h\mathbf{u}) - f(\mathbf{x})}{h}.$$

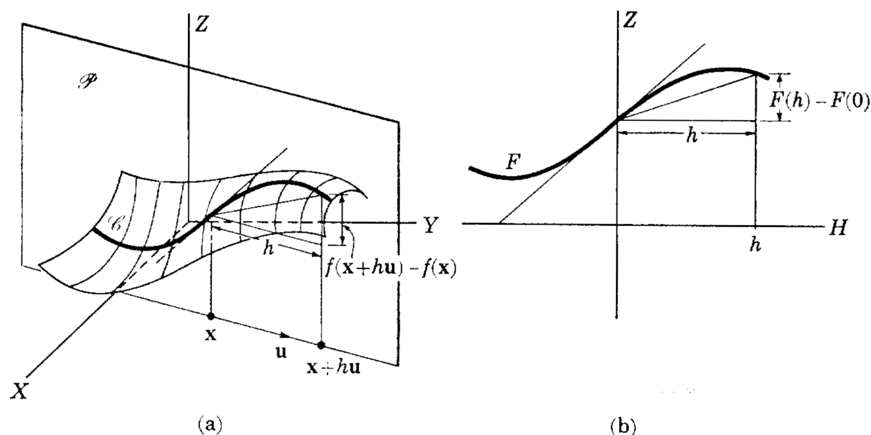


FIGURA 13

Esta razón de cambio se llama *derivada direccional* de f en la dirección $\mathbf{u}$ en el punto $\mathbf{x}$ y se representa por el símbolo $D_{\mathbf{u}}f(\mathbf{x})$.

Nótese que el valor de la derivada direccional $D_{\mathbf{u}}f(\mathbf{x})$ depende solamente de los valores de la función f en los puntos sobre la recta que pasa por $\mathbf{x}$ y es paralela a $\mathbf{u}$: $\{\mathbf{x}+h\mathbf{u} \mid h \in \mathrm{R}\}$. Los puntos sobre esta recta quedan especificados por el valor del parámetro h. Si definimos la función F por la regla

$$F(h) = f(\mathbf{x}+h\mathbf{u}),$$

entonces F es una función real de una variable real cuya gráfica corresponde a la intersección de la gráfica de f con el plano $\mathscr{P}$ que es perpendicular al plano XY y contiene la recta que pasa por $\mathbf{x}$ y es paralela a $\mathbf{u}$ (figura 13). El punto $(0, F(0))$ corresponde al punto $(x, y, f(x, y))$. Como

7.2 $$D_{\mathbf{u}}f(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{u})-f(\mathbf{x})}{h} = \lim_{h\to 0} \frac{F(h)-F(0)}{h} = F'(0),$$

decimos que $D_{\mathbf{u}}f(\mathbf{x})$ es la pendiente en $(x, y, f(x, y))$ de la curva $\mathscr{C}$ formada por la intersección de la gráfica de f y el plano $\mathscr{P}$.

Definimos ahora la derivada direccional en el caso general de una función de R^n a R. En esta definición suponemos que la función está definida sobre un conjunto abierto en R^n.

7.3 Definición. *La* ***derivada direccional de*** f ***en la dirección*** $\mathbf{u}$, *donde* $\mathbf{u}$ *es un vector unitario en* R^n, *es la función* $D_{\mathbf{u}}f$ *con regla de correspondencia*

$$D_{\mathbf{u}}f(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{u})-f(\mathbf{x})}{h}$$

y con dominio el conjunto de todos los puntos $\mathbf{x}$ *del dominio de* f *para los que existe tal límite.*

Nota. Obsérvese que si f es una función de R en R y $\mathbf{u}$ se hace igual al número 1, entonces esta definición es la misma que la definición de la derivada de una función real de variable real.

El valor de la derivada direccional, $D_{\mathbf{u}}f(\mathbf{x})$, es la razón de cambio de f en la dirección $\mathbf{u}$ en el punto $\mathbf{x}$. Por ejemplo, si $f(x,y,z)$ es la temperatura en el punto (x, y, z), entonces $D_{\mathbf{u}}f(x, y, z)$ es la razón del cambio de temperatura en (x, y, z) con respecto a la distancia a lo largo de la recta que pasa por (x, y, z) en la dirección $\mathbf{u}$. Consideremos como otro ejemplo la función f de R^4 en R donde $f(x, y, z, t)$ es la temperatura en el punto (x, y, z) en el instante t. Si $\mathbf{u} = (u_1, u_2, u_3, 0)$, entonces $D_{\mathbf{u}}f(x, y, z, t)$ es la razón de cambio de la temperatura [en el punto (x, y, z) y en el instante t] con respecto a la distancia a lo largo de la recta que pasa por (x, y, z) en la dirección (u_1, u_2, u_3). Y, si $\mathbf{u} = (0, 0, 0, 1)$, entonces $D_{\mathbf{u}}f(x,y,z,t)$ es la

razón de cambio de la temperatura [en el punto (x, y, z) en el instante t] con respecto al tiempo.

Si para $\mathbf{x}$ y $\mathbf{u}$ fijos hacemos $F(h) = f(\mathbf{x}+h\mathbf{u})$, entonces, como antes vimos

$$\textbf{7.4} \qquad D_{\mathbf{u}}f(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{u})-f(\mathbf{x})}{h} = \lim_{h\to 0} \frac{F(h)-F(0)}{h} = F'(0).$$

Así pues, nada nuevo necesitamos aprender para calcular derivadas direccionales: pueden calcularse diferenciando una función real de variable real. Usamos esto en la solución 2 del siguiente ejemplo.

7.5 Ejemplo. Determínese $D_{\mathbf{u}}f$ donde $f = I_1^2 + I_2^2 + I_3^2$ y $\mathbf{u} = \frac{1}{\sqrt{6}}(2, -1, 1)$.

Solución 1.

$$\lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{u})-f(\mathbf{x})}{h}$$

$$= \lim_{h\to 0} \frac{1}{h}\left[\left(x + \frac{2}{\sqrt{6}}h\right)^2 + \left(y - \frac{1}{\sqrt{6}}h\right)^2 + \left(z + \frac{1}{\sqrt{6}}h\right)^2 - x^2 - y^2 - z^2\right]$$

$$= \lim_{h\to 0} \left(\frac{4}{\sqrt{6}}x + \frac{4}{6}h - \frac{2}{\sqrt{6}}y + \frac{1}{6}h + \frac{2}{\sqrt{6}}z + \frac{1}{6}h\right)$$

$$= \frac{2}{\sqrt{6}}(2x - y + z).$$

Por tanto,

$$D_{\mathbf{u}}f(x, y, z) = \frac{2}{\sqrt{6}}(2x - y + z)$$

y

$$D_{\mathbf{u}}f = \frac{2}{\sqrt{6}}(2I_1 - I_2 + I_3).$$

Solución 2. Si hacemos $F(h) = f(\mathbf{x}+h\mathbf{u})$ entonces $D_{\mathbf{u}}f(\mathbf{x}) = F'(0)$.

$$F(h) = f(\mathbf{x}+h\mathbf{u}) = \left(x + \frac{2}{\sqrt{6}}h\right)^2 + \left(y - \frac{1}{\sqrt{6}}h\right)^2 + \left(z + \frac{1}{\sqrt{6}}h\right)^2$$

y

$$F'(h) = \frac{4}{\sqrt{6}}\left(x + \frac{2}{\sqrt{6}}h\right) - \frac{2}{\sqrt{6}}\left(y - \frac{1}{\sqrt{6}}h\right) + \frac{2}{\sqrt{6}}\left(z + \frac{1}{\sqrt{6}}h\right).$$

Luego

$$D_{\mathbf{u}}f(\mathbf{x}) = F'(0) = \frac{4}{\sqrt{6}}x - \frac{2}{\sqrt{6}}y + \frac{2}{\sqrt{6}}z$$

y

$$D_{\mathbf{u}}f = \frac{2}{\sqrt{6}}(2I_1 - I_2 + I_3).$$

Aunque la derivada direccional parece una extensión natural de la derivada de una función real de variable real, hay algunas propiedades importantes de ésta que no se conservan en esta extensión. Por ejemplo, una función puede tener en un punto derivadas direccionales en todas las direcciones y, sin embargo, no ser continua en ese punto.

7.6 Ejemplo. Sea f la función de R^2 en R definida por

$$f(x, y) = \frac{x^2 y}{y^2 + x^4}, \quad (x, y) \neq (0, 0)$$

$$f(0, 0) = 0.$$

Demuéstrese que en el origen existe la derivada direccional de f en cualquier dirección, pero que f no es continua en el origen.

SOLUCIÓN. Sea $\mathbf{u} = (u_1, u_2)$ un vector unitario. Entonces

$$D_{\mathbf{u}}f(0, 0) = \lim_{h \to 0} \frac{f(hu_1, hu_2) - f(0, 0)}{h}$$

$$= \lim_{h \to 0} \frac{1}{h}\left[\frac{h^3 u_1^2 u_2}{h^2 u_2^2 + h^4 u_1^4}\right]$$

$$= \lim_{h \to 0} \frac{u_1^2 u_2}{u_2^2 + h^2 u_1^4}$$

$$= \begin{cases} \dfrac{u_1^2}{u_2}, & \text{si } u_2 \neq 0 \\ 0, & \text{si } u_2 = 0. \end{cases}$$

Así pues, en el origen, la derivada direccional de f en cualquier dirección existe. Para demostrar que f no es continua en el origen, consideremos la parábola $\mathscr{E} = \{(x, y) \mid y = x^2\}$. Entonces

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y}{y^2+x^4} = \lim_{(x,y)\to(0,0)} \tfrac{1}{2} = \tfrac{1}{2} \quad (\text{sobre } \mathscr{E}).$$

Como $f(0, 0) = 0$, f no es continua en $(0, 0)$.

Ahora demostraremos que si f es diferenciable en un punto, entonces todas las derivadas direccionales existen en ese punto.

7.7 Teorema. *Si f es diferenciable en* $\mathbf{x}$, *entonces la derivada direccional de f en cualquier dirección* $\mathbf{u}$ *existe en* $\mathbf{x}$ *y*

$$D_{\mathbf{u}}f(\mathbf{x}) = \mathbf{D}f(\mathbf{x}) \cdot \mathbf{u} = df(\mathbf{x}; \mathbf{u}).$$

PRUEBA. Si hacemos $g(h) = \mathbf{x}+h\mathbf{u}$ y $F = f \circ \mathbf{g}$, entonces $F(h) = f(\mathbf{x}+h\mathbf{u})$ y de acuerdo con 7.4, $D_{\mathbf{u}}f(\mathbf{x}) = F'(0)$. Usando el teorema 6.9 obtenemos

$$F'(0) = D[f \circ \mathbf{g}](0) = \mathbf{D}f(\mathbf{g}(0))\text{-}D\mathbf{g}(0)$$

y, por tanto,

$$D_{\mathbf{u}}f(\mathbf{x}) = \mathbf{D}f(\mathbf{x}) \cdot \mathbf{u} = df(\mathbf{x}; \mathbf{u}).$$

El recíproco de este teorema no se verifica. Una función de R^n en R puede tener derivadas direccionales en todas las direcciones en un punto y no ser diferenciable en ese punto. La función considerada en el ejemplo 7.6 tiene derivadas direccionales en todas direcciones en el origen, pero no es ni tan siquiera continua en el origen, y, por tanto, ciertamente no diferenciable.

Como el producto escalar de dos vectores alcanza su valor máximo cuando los vectores están en la misma dirección, $D_{\mathbf{u}}f(\mathbf{x}) = \mathbf{D}f(\mathbf{x}) \cdot \mathbf{u} = \text{Comp}_{\mathbf{u}}\, \mathbf{D}f(\mathbf{x})$ implica que $\mathbf{D}f(\mathbf{x})$ es un vector en la dirección de la máxima razón de cambio de f y que esta máxima razón de cambio es la longitud de $\mathbf{D}f(\mathbf{x})$.

Nótese que si hacemos $\mathbf{h} = h\mathbf{u}$ donde $h = |\mathbf{h}|$, entonces

$$df(\mathbf{x}; \mathbf{h}) = h\mathbf{u} \cdot \mathbf{D}f(\mathbf{x}) = hD_{\mathbf{u}}f(\mathbf{x}).$$

Así pues, el valor $df(\mathbf{x}; \mathbf{h})$ de la diferencial es la longitud de $\mathbf{h}$ veces el valor en $\mathbf{x}$ de la derivada direccional de f en la dirección $\mathbf{u} = \dfrac{\mathbf{h}}{|\mathbf{h}|}$.

Usando los teoremas 7.7, 6.6 y 6.7 y el corolario 6.8, podíamos proceder a derivar las fórmulas para la derivada direccional de la suma, producto, composición y cociente de funciones diferenciables (problemas 6, 7 y 8). Sin embargo, estas fórmulas no son necesarias para propósitos de cálculo.

En la sección 9 derivaremos un método sencillo para obtener las derivadas direccionales.

Para derivadas direccionales podemos probar un teorema del valor medio que es análogo al dado para las derivadas (teorema 6.10).

7.8 Teorema. (Teorema del valor medio.) *Si $D_{\mathbf{u}}f$ existe sobre un conjunto abierto que contiene el segmento rectilíneo cerrado que va de* $\mathbf{x}$ *a* $\mathbf{x}+h\mathbf{u}$, *donde* $\mathbf{u}$ *es un vector unitario, entonces existe un número $\theta \in \langle 0, 1 \rangle$ tal que*

$$f(\mathbf{x}+h\mathbf{u})-f(\mathbf{x}) = hD_{\mathbf{u}}f(\mathbf{x}+\theta h\mathbf{u}).$$

PRUEBA. Si hacemos $F(t) = f(\mathbf{x}+t\mathbf{u})$, $t \in [0, h]$, entonces, usando la definición de derivada direccional, vemos que $F'(t) = D_{\mathbf{u}}f(\mathbf{x}+t\mathbf{u})$. De donde F resulta ser continua sobre $[0, h]$ y diferenciable sobre $\langle 0, h \rangle$. Aplicando a F el teorema del valor medio para funciones de R en R, obtenemos

$$F(h)-F(0) = hF'(\theta h) \quad \text{para algún} \quad \theta \in \langle 0, 1 \rangle.$$

Luego,

$$f(\mathbf{x}+h\mathbf{u})-f(\mathbf{x}) = hD_{\mathbf{u}}f(\mathbf{x}+\theta h\mathbf{u}).$$

Problemas

1. Usando la definición de $D_{\mathbf{u}}f$, determínese $D_{\mathbf{u}}f$ cuando

a) $f(x, y) = x^2 y$, $\mathbf{u} = \mathbf{i} = (1, 0)$
b) $f(x, y) = x^2 y$, $\mathbf{u} = \mathbf{j} = (0, 1)$
c) $f(x, y, z) = xz+y$, $\mathbf{u} = \dfrac{1}{\sqrt{42}}(-5, 4, 1)$
d) $f = I_1 I_2{}^2 + I_3$, $\mathbf{u} = \dfrac{1}{\sqrt{6}}(1, 2, 1)$.

2. Si $f(x, y, z) = \dfrac{xz+4}{x+y}$, determínese $D_{\mathbf{u}}f$ cuando

a) $\mathbf{u} = \mathbf{i} = (1, 0, 0)$ *b*) $\mathbf{u} = \mathbf{j} = (0, 1, 0)$ *c*) $\mathbf{u} = \mathbf{k} = (0, 0, 1)$.

3. Encuéntrese ${}_{\mathbf{u}}Df$ mediante la determinación de $F'(0)$, donde $F(h) = f(\mathbf{x}+h\mathbf{u})$, cuando

a) $f(x, y) = \operatorname{sen} xy$, $\mathbf{u} = \dfrac{1}{\sqrt{13}}(-2, 3)$
b) $f(x, y, z) = e^{xy} + \ln z$, $\mathbf{u} = \dfrac{1}{\sqrt{13}}(2, 0, 3)$.

4. Demuéstrese que $D_{\mathbf{u}}c = 0$ cuando c es una constante de R^n en R y $\mathbf{u}$ un vector unitario en R^n.

5. Demuéstrese que $D_{-\mathbf{u}}f = -D_{\mathbf{u}}f$.

6. Si f y g son diferenciables sobre un conjunto abierto $\mathscr{E}$, demuéstrese que sobre $\mathscr{E}$

$$D_{\mathbf{u}}(f+g) = D_{\mathbf{u}}f + D_{\mathbf{u}}g$$

$$D_{\mathbf{u}}(fg) = fD_{\mathbf{u}}g + gD_{\mathbf{u}}f.$$

7. Si f es una función de R^n en R que es diferenciable sobre un conjunto abierto $\mathscr{E}$ y g es una función de R en R que es diferenciable sobre un conjunto abierto que contiene a $f(\mathscr{E})$, demuéstrese que sobre $\mathscr{E}$

$$D_{\mathbf{u}}(g \circ f) = (g' \circ f)D_{\mathbf{u}}f.$$

8. Si f y g son diferenciables sobre un conjunto abierto $\mathscr{E}$ y g es distinta de cero sobre $\mathscr{E}$, demuéstrese que sobre $\mathscr{E}$

$$D_{\mathbf{u}}\left(\frac{f}{g}\right) = \frac{gD_{\mathbf{u}}f - fD_{\mathbf{u}}g}{g^2}.$$

9. Si f es diferenciable sobre $\mathbf{x}_0$ y f tiene un máximo relativo en $\mathbf{x}_0$, demuéstrese que $D_{\mathbf{u}}f(\mathbf{x}_0) = 0$ para toda dirección $\mathbf{u}$.

10. Si f es una función de R^2 en R que es diferenciable en el punto $\mathbf{x}$, demuéstrese que

$$\mathbf{D}f(\mathbf{x}) = (D_{\mathbf{i}}f(\mathbf{x}), D_{\mathbf{j}}f(\mathbf{x})), \quad \text{donde} \quad \mathbf{i} = (1, 0) \text{ y } j = (0, 1).$$

11. Un conjunto $\mathscr{E}$ se llama *convexo* si, para cualesquier dos puntos $\mathbf{x}$ y $\mathbf{y}$ en $\mathscr{E}$, el segmento rectilíneo de $\mathbf{x}$ a $\mathbf{y}$ se encuentra en $\mathscr{E}$. Pruébese que si para todo $\mathbf{u}$, $D_{\mathbf{u}}f(\mathbf{x}) = 0$ en todo punto $\mathbf{x}$ de un conjunto abierto y convexo $\mathscr{E}$, entonces f es constante sobre $\mathscr{E}$.

8. DERIVADAS PARCIALES

Las derivadas direccionales en la dirección de los ejes de coordenadas se llaman derivadas parciales.

8.1 Definición. *Si f es una función real definida sobre un conjunto abierto de R^n y $\mathbf{u}_k$ es el vector con componente k-ésimo 1 y todos los otros componentes 0, entonces llamamos a $D_{\mathbf{u}_k}f$ la* ***derivada parcial*** *de f con respecto a la k-ésima coordenada.*

Por brevedad denotaremos $D_{\mathbf{u}_k}f$ por D_kf. Así pues, D_kf es la función con regla de correspondencia,

8.2 $$D_kf(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{u}_k)-f(\mathbf{x})}{h}$$

y dominio el conjunto de puntos $\mathbf{x}$ en el dominio de f para los que el límite que aparece en 8.2 existe.

Sea f una función de R^2 en R. Supongamos que la gráfica de f,

$$\{(x, y, z) \mid z = f(x, y)\},$$

es la superficie dibujada en la figura 14a. De acuerdo con 8.2, las derivadas parciales de f en el punto $(\mathbf{x}_0, \mathbf{y}_0)$ vienen dadas por

$$D_1f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0+h, y_0)-f(x_0, y_0)}{h}$$

$$D_2f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0, y_0+h)-f(x_0, y_0)}{h}.$$

Estas derivadas parciales pueden calcularse por el método dado para calcular derivadas direccionales o por un método ligeramente diferente que es especialmente adecuado para el cálculo de las derivadas parciales. Claramente, el valor de $D_1f(x_o,y_o)$ depende solamente de los valores de f en puntos sobre la recta paralela al eje X que pasa por el punto (x_o,y_o). Los puntos sobre esta recta pueden describirse simplemente dando solamente su primera coordenada; es decir, x especifica el punto (x, y_0) sobre la recta $y = y_0$. Si definimos la función g_1 por la regla

$$g_1(x) = f(x, y_0),$$

entonces

$$D_1f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0+h, y_0)-f(x_0, y_0)}{h}$$

$$= \lim_{h\to 0} \frac{g_1(x_0+h)-g_1(x_0)}{h} = g_1'(x_0).$$

Análogamente, si definimos g_2 por

$$g_2(y) = f(x_0, y),$$

entonces

$$D_2f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0, y_0+h)-f(x_0, y_0)}{h}$$

$$= \lim_{h\to 0} \frac{g_2(y_0+h)-g_2(y_0)}{h} = g_2'(y)_0\mathbf{i}.$$

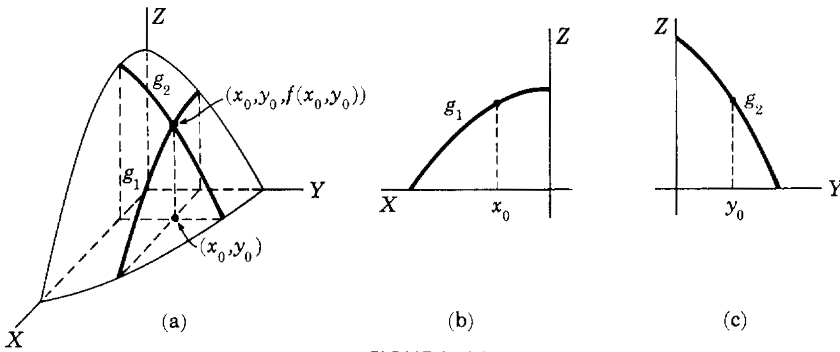


FIGURA 14

Las gráficas de g_1 y g_2 (figuras 14b y c) son la curvas de intersección de la superficie $z = f(x, y)$ con los planos $y = y_0$ y $x = x_0$, respectivamente.

8.3 Ejemplo. Si $f = I_1{}^2 + I_2{}^2$, determínense las derivadas parciales de f.

SOLUCIÓN. Si para cualquier punto $(x_0, y_0) \in R^2$ (figura 15), hacemos $g_1(x) = f(x, y_0)$, entonces $D_1 f(x_0, y_0) = g_1'(x_0)$. Como $f = I_1{}^2 + I_2{}^2$,

$$g_1(x) = x^2 + y_0{}^2$$

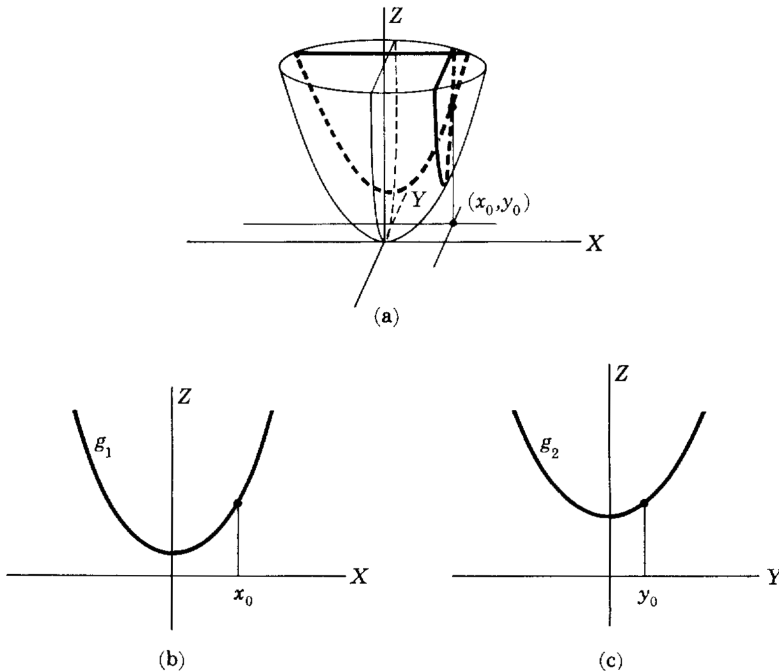


FIGURA 15

y

$$g_1'(x) = 2x.$$

Así pues, $D_1 f(x_0, y_0) = 2x_0$ y $D_1 f = 2I_1$. Haciendo $_2(y) = f(x_0, y)$, tenemos $D_2 f(x_0, y_0) = g_2'(y_0)$. Luego

$$g_2(y) = x_0^2 + y^2$$

y

$$g_2'(y) = 2y.$$

Así pues, $D_2 f(x_0, y_0) = 2y_0$ y $D_2 f = 2I_2$.

En la resolución del ejemplo 8.3 no necesitamos introducir explícitamente las funciones g_1 y g_2. Para encontrar $D_1 f$ consideramos y en

8.4 $$f(x, y) = x^2 + y^2$$

como número fijo. Entonces, para un y fijo, 8.4 define una función de una variable real (la función g_1) cuya derivada nos da $D_1 f$. Análogamente, para encontrar $D_2 f$, se considera x en 8.4 como número fijo. Entonces, para un x fijo, 8.4 define una función de una variable real (la función g_2) cuya derivada nos da $D_2 f$.

Este método puede emplearse en el caso general para encontrar las derivadas parciales de una función de R^n en R. Para encontrar la derivada parcial de f respecto a la k-ésima coordenada, consideramos todas las coordenadas excepto la k-ésima como números fijos. Obtenemos entonces una función de una variable real cuya derivada es $D_k f$. Es decir, si

$$g_k(x_k) = f(a_1, \ldots, x_k, \ldots, a_n)$$

entonces

$$D_k f(\mathbf{a}) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_k + h, \ldots, a_n) - f(a_1, \ldots, a_k, \ldots, a_n)}{h}$$

$$= \lim_{h \to 0} \frac{g_k(a_k + h) - g_k(a_k)}{h} = g_k'(a_k).$$

Así pues, el problema de encontrar las derivadas parciales, lo mismo que el de encontrar las derivadas direccionales de cualquier dirección, se reduce al de diferenciar una función real de una variable real.

8.5 Ejemplo. Encuéntrense las derivadas parciales de la función f definida por $f(x, y, z) = xy + \cos(yz)$.

SOLUCIÓN. Considerando y y z como números fijos, diferenciamos la función

$$g_1(x) = f(x, y, z) = xy + \cos(yz).$$

Entonces,

$$D_1 f(x, y, z) = y.$$

Considerando x y z como números fijos, obtenemos

$$D_2 f(x, y, z) = x - z \operatorname{sen}(yz).$$

Considerando x y y como números fijos, obtenemos

$$D_3 f(x, y, z) = -y \operatorname{sen}(yz).$$

Las notaciones para derivadas parciales son muchas y variadas. Si f es una función de R^2 en R y si hacemos $z = f(x, y)$, las derivadas parciales pueden denotarse como sigue:

$$D_1 f(x, y) = f_1(x, y) = f_x(x, y) = \frac{\partial f(x, y)}{\partial x} = \frac{\partial z}{\partial x}$$

$$D_2 f(x, y) = f_2(x, y) = f_y(x, y) = \frac{\partial f(x, y)}{\partial y} = \frac{\partial z}{\partial y}.$$

Por ejemplo, si $f = I_1^2 + I_2^2$ (ejemplo 8.3) y si hacemos

$$z = f(x, y) = x^2 + y^2,$$

entonces las derivadas parciales pueden denotarse por

$$\frac{\partial z}{\partial x} = 2x \quad \text{y} \quad \frac{\partial z}{\partial y} = 2y.$$

Si $f = I_1 I_2 + \cos \circ (I_2 I_3)$ (ejemplo 8.5) y si hacemos

$$w = f(x, y, z) = xy + \cos(yz),$$

entonces las derivadas parciales pueden denotarse por

$$\frac{\partial w}{\partial x} = y, \quad \frac{\partial w}{\partial y} = x - z \operatorname{sen}(yz), \quad \frac{\partial w}{\partial z} = -y \operatorname{sen}(yz).$$

La consideración de las derivadas parciales de una función puede proveer una forma sencilla de demostrar que la función es diferenciable. El ejemplo 7.6 muestra que la existencia de todas las derivadas parciales no es suficiente para garantizar que la función es diferenciable. Sin embargo, ahora vamos a probar que la continuidad de las derivadas parciales nos permite afirmar que la función es diferenciable.

8.6 Teorema. *Si todas las derivadas parciales de f existen y son continuas sobre un conjunto abierto $\mathscr{E}$, entonces f es diferenciable sobre $\mathscr{E}$.*

PRUEBA. Damos la prueba para una función de R^2 en R. Sin embargo, el método empleado es general y con la introducción de una notación adecuada podría adaptarse para una función de R^n en R.

Tomemos $\mathbf{x} = (x, y) \in \mathscr{E}$ y sea $\mathscr{S}(\mathbf{x}; r)$ una vecindad de $\mathbf{x}$ contenida en $\mathscr{E}$. Tomemos $\mathbf{h} = (h_1, h_2)$ tal que $|\mathbf{h}| < r$. Entonces, usando el teorema del valor medio (teorema 7.8, pág. 200), tenemos

$$\begin{aligned} f(\mathbf{x}+\mathbf{h})-f(\mathbf{x}) &= f(x+h_1, y+h_2)-f(x, y) \\ &= f(x+h_1, y+h_2)-f(x, y+h_2)+f(x, y+h_2)-f(x, y) \\ &= h_1 D_1 f(x+\theta_1 h_1, y+h_2)+h_2 D_2 f(x, y+\theta_2 h_2) \end{aligned}$$

donde $\theta_i \in \langle 0, 1 \rangle$, $i = 1, 2$. Como $D_1 f$ y $D_2 f$ son continuas sobre $\mathscr{E}$,

$$\begin{aligned} D_1 f(x+\theta_1 h_1, y+h_2) &= D_1 f(x, y)+\varphi_1(\mathbf{x}; \mathbf{h}) \\ D_2 f(x, y+\theta_2 h_2) &= D_2 f(x, y)+\varphi_2(\mathbf{x}; \mathbf{h}) \end{aligned}$$

donde $\lim_{\mathbf{h}\to\mathbf{0}} \varphi_i(\mathbf{x}; \mathbf{h}) = 0$, $i = 1, 2$.

Luego,

$$\begin{aligned} f(\mathbf{x}+\mathbf{h})-f(\mathbf{x}) &= h_1(D_1 f(x, y)+\varphi_1(\mathbf{x}; \mathbf{h}))+h_2(D_2 f(x, y)+\varphi_2(\mathbf{x}; \mathbf{h})) \\ &= (D_1 f(\mathbf{x}), D_2 f(\mathbf{x}))\cdot\mathbf{h}+\boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})\cdot\mathbf{h}, \end{aligned}$$

donde $\boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) = (\varphi_1(\mathbf{x}; \mathbf{h}), \varphi_2(\mathbf{x}; \mathbf{h}))$. Como $\lim_{\mathbf{h}\to\mathbf{0}} \varphi_i(\mathbf{x}; \mathbf{h}) = 0$, $i = 1, 2$, $\lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) = \mathbf{0}$ y esto nos dice que f es diferenciable en $\mathbf{x}$.

El recíproco de este teorema no se verifica (problema 4).

Como las derivadas parciales de un polinomio son ellas mismas polinomios y, por tanto, continuas en todos los puntos de R^n, el teorema 8.6 implica que los polinomios son diferenciables sobre R^n. Análogamente, las funciones racionales son diferenciables en todos los puntos en que están definidas.

Problemas

1. Determínense las derivadas parciales de f cuando

a) $f(x, y) = x^2 y^2 + y$ *b*) $f(x, y) = \dfrac{xy}{x-y}$

c) $f(x, y, z) = \dfrac{x^2 z + y + 2}{x^2 + 1}$ *d*) $f(x, y, z) = \text{sen}\,(xy+z)$

e) $f(x, y) = \arctan\left(\dfrac{y}{x}\right)$ *f*) $f(x, y, z) = x^2 z + e^{yz}$

g) $f(x, y) = \sqrt{x^2 + y^2}$ *h*) $f(x, y, z) = z \ln(x+y)$

2. Si las derivadas parciales de f y g con respecto a la k-ésima coordenada existen sobre un conjunto abierto $\mathscr{E}$, demuéstrese que sobre $\mathscr{E}$,

a) $D_k(f+g) = D_kf + D_kg$
b) $D_k(fg) = fD_kg + gD_kf$
c) $D_k\left(\dfrac{f}{g}\right) = \dfrac{gD_kf - fD_kg}{g^2}$ si g es distinto de cero sobre $\mathscr{E}$

3. Si I_j es la j-ésima función proyección $(I_j(\mathbf{x}) = x_j)$ pruébese que

$$D_kI_j = \delta_{kj} \quad \text{donde} \quad \delta_{kj} = \begin{cases} 0 & \text{si } k \neq j \\ 1 & \text{si } k = j. \end{cases}$$

4. Demuéstrese que la función f definida por

$$f(x, y) = (x^2+y^2)\,\text{sen}\,\frac{1}{\sqrt{x^2+y^2}}, \quad (x, y) \neq (0, 0)$$

$$f(0, 0) = 0$$

es diferenciable en el origen, pero no tiene derivadas parciales continuas en el origen.

9. ALGUNOS EJEMPLOS

En esta sección discutiremos técnicas para determinar la diferencial, la derivada y las derivadas direccionales de funciones diferenciables de R^n en R. El teorema 7.7, pág. 199, nos dice que si f es diferenciable en $\mathbf{x}$ entonces

9.1 $$D_{\mathbf{u}}f(\mathbf{x}) = \mathbf{D}f(\mathbf{x}) \cdot \mathbf{u}$$

donde $\mathbf{u}$ es un vector unitario. Si $\mathbf{u}_k$ denota el vector unitario con 1 como componente k-ésima y todas las demás componentes iguales a cero, entonces

$$D_kf(\mathbf{x}) = \mathbf{D}f(\mathbf{x}) \cdot \mathbf{u}_k.$$

Como $\mathbf{D}f(\mathbf{x}) \cdot \mathbf{u}_k$ es la k-ésima componente de $\mathbf{D}f(\mathbf{x})$, tenemos

9.2 $$\mathbf{D}f(\mathbf{x}) = (D_1f(\mathbf{x}), \ldots, D_nf(\mathbf{x})).$$

El vector $(D_1f(\mathbf{x}), \ldots, D_nf(\mathbf{x}))$ se llama *gradiente* de f en $\mathbf{x}$ y se denota por $\nabla f(\mathbf{x})$ (y se lee "nabla f en $\mathbf{x}$" o "gradiente de f en $\mathbf{x}$"). Hemos, pues, demostrado que si f es diferenciable en $\mathbf{x}$, entonces la derivada de f en $\mathbf{x}$ es el gradiente de f en $\mathbf{x}$.

Como las derivadas parciales son fáciles de calcular, 9.2 nos facilita un método conveniente para determinar los valores de la derivada y la

diferencial. Luego, usando 9.1, podemos encontrar los valores de las derivadas direccionales. Damos a continuación algunos ejemplos para ilustrar estos métodos.

Primero, una palabra acerca de la notación. Al discutir las diferenciales es práctica común usar $d\mathbf{x}$ en lugar de $\mathbf{h}$. Además, si f es una función de R^3 en R y si $w = f(x, y, z)$, escribimos $df(\mathbf{x}; d\mathbf{x})$ como sigue:

$$dw = \nabla f(\mathbf{x})\cdot d\mathbf{x} = \left(\frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial w}{\partial z}\right)\cdot(dx, dy, dz)$$

$$= \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial y}dy + \frac{\partial w}{\partial z}dz.$$

9.3 Ejemplo. Si f es la función de R^3 en R definida por $f(x, y, z) = x^2y + xe^z$, determínese la diferencial de f.

SOLUCIÓN. El gradiente de f está dado por

$$\nabla f(x, y, z) = (2xy + e^z, x^2, xe^z).$$

Como el gradiente de f es continuo sobre R^3, el teorema 8.6 implica que f es diferenciable sobre R^3. Haciendo $w = f(x, y, z)$, tenemos

$$df(\mathbf{x}; d\mathbf{x}) = dw = (2xy + e^z, x^2, xe^z)\cdot(dx, dy, dz)$$

$$= (2xy + e^z)dx + x^2\,dy + xe^z\,dz.$$

9.4 Ejemplo. Determínese $D_{\mathbf{u}}f$ donde $f = I_1{}^2 + I_1 I_2$ y $\mathbf{u} = \frac{1}{\sqrt{5}}(2, -1)$.

SOLUCIÓN. El gradiente de f es $\nabla f = (2I_1 + I_2, I_1)$. Como el gradiente es continuo, f es diferenciable sobre R^2 y, por tanto, $D_{\mathbf{u}}f = \mathbf{u}\cdot\nabla f$. Por tanto,

$$D_{\mathbf{u}}f = \frac{1}{\sqrt{5}}(2, -1)\cdot(2I_1 + I_2, I_1) = \frac{1}{\sqrt{5}}(3I_1 + 2I_2).$$

9.5 Ejemplo. Encuéntrese la dirección y magnitud de la razón de cambio máxima de la función $f = I_1{}^2 + I_1 I_2$ en el punto (2, 3).

SOLUCIÓN. El gradiente es un vector que tienen la dirección y magnitud de la razón de cambio máxima de la función. Como $\nabla f(x, y) = (2x + y, x)$, $\Delta f(2, 3) = (7, 2)$. Así pues, f tiene su razón de cambio máxima en (2, 3) en la dirección $\mathbf{u} = \frac{1}{\sqrt{53}}(7, 2)$ y la magnitud de esta razón de cambio máxima es

$$|\nabla f(2, 3)| = |(7, 2)| = \sqrt{53}.$$

Supongamos $w = f(\mathbf{x})$, donde f es una función real que es diferenciable sobre R^3, y supongamos que $\mathscr{C}$ es una curva lisa en R^3 descrita por la ecuación $\mathbf{x} = \mathbf{h}(t)$, $t \in \mathscr{J}$. Consideremos el problema de encontrar la razón de cambio de f con respecto a la distancia a lo largo de la curva $\mathscr{C}$. Si hacemos $s = l(t) = \int_{t_0}^{t} |\mathbf{h}'|$, donde $t_0 \in \mathscr{J}$, entonces s es la longitud del arco de $\mathscr{C}$ desde el punto $\mathbf{h}(t_0)$ hasta el punto $\mathbf{h}(t)$. Como $\mathscr{C}$ es una curva lisa, l es una función creciente y, por tanto, tiene una inversa l^*. Haciendo $\mathbf{g} = \mathbf{h} \circ l^*$, podemos escribir $\mathbf{x} = \mathbf{g}(s)$ y $w = f(\mathbf{g}(s))$ sobre $\mathscr{C}$. Entonces la razón de cambio de f con respecto a la distancia a lo largo de la curva es $\dfrac{dw}{ds}$.

Usando el teorema 6.9, pág. 193, tenemos

$$D[f \circ \mathbf{g}](s) = \mathbf{D}f(\mathbf{g}(s)) \cdot D\mathbf{g}(s),$$

es decir,

$$\frac{dw}{ds} = \left(\frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial w}{\partial z}\right) \cdot \frac{d\mathbf{x}}{ds}.$$

Según el teorema 7.7, pág. 126, y 11.4′, pág. 144,

$$\frac{d\mathbf{x}}{ds} = \frac{d\mathbf{x}}{dt}\frac{dt}{ds} = \frac{ds}{dt}\mathbf{T}\frac{dt}{ds} = \mathbf{T}$$

donde $\mathbf{T}$ es el vector unitario tangente a $\mathscr{C}$. Así pues, la razón de cambio de f con respecto a la distancia a lo largo de $\mathscr{C}$ es la derivada direccional de f en la dirección $\mathbf{T}$: $\dfrac{dw}{ds} = D_{\mathbf{T}}f$.

9.6 Ejemplo. Si $f(x, y, z) = x^2 + y^2 + z^2$ y $\mathscr{C}$ es la hélice cilíndrica definida por

$$(x, y, z) = (\cos t, \operatorname{sen} t, \tfrac{1}{2}t), \quad t \in \langle -\infty, \infty \rangle,$$

encuéntrese la razón de cambio de f con respecto a la distancia a lo largo de la curva $\mathscr{C}$ en el punto $\left(0, 1, \dfrac{\pi}{4}\right)$.

SOLUCIÓN 1. Haciendo $w = f(x, y, z)$, expresamos w en términos de s y encontramos $\dfrac{dw}{ds}$. Sea $\mathbf{g}(t) = (\cos t, \operatorname{sen} t, \tfrac{1}{2}t)$. Entonces

$$\mathbf{g}'(t) = (-\operatorname{sen} t, \cos t, \tfrac{1}{2})$$

y

$$s = \int_0^t |\mathbf{g}'| = \int_0^t |(-\text{sen}, \cos, \tfrac{1}{2})| = \int_0^t \frac{\sqrt{5}}{2} = \frac{\sqrt{5}}{2}t.$$

Por tanto, sobre $\mathscr{C}$

$$w = \cos^2 t + \text{sen}^2 t + \tfrac{1}{4}t^2 = 1 + \tfrac{1}{4}t^2 = 1 + \tfrac{1}{5}s^2,$$

y la razón de cambio de w con respecto a s es

$$\frac{dw}{ds} = \frac{2}{5}s.$$

Como el punto $\left(0, 1, \frac{\pi}{4}\right)$ corresponde a $t = \frac{\pi}{2}$ y a $s = \frac{\sqrt{5}}{4}\pi$, la razón de cambio de f con respecto a la distancia a lo largo de $\mathscr{C}$ en $\left(0, 1, \frac{\pi}{4}\right)$ es $\frac{\sqrt{5}}{10}\pi$.

SOLUCIÓN 2. En esta solución usamos el hecho de que la razón de cambio de f con respecto a la distancia a lo largo de $\mathscr{C}$ en $\mathbf{x}$ es $D_{\mathbf{T}}f(\mathbf{x})$. En $\left(0, 1, \frac{\pi}{4}\right)$, el vector unitario tangente es $\mathbf{T} = \frac{2}{\sqrt{5}}\left(-1, 0, \frac{1}{2}\right)$. Entonces

$$D_{\mathbf{T}}f\left(0, 1, \frac{\pi}{4}\right) = \nabla f\left(0, 1, \frac{\pi}{4}\right) \cdot \frac{2}{\sqrt{5}}\left(-1, 0, \frac{1}{2}\right)$$

$$= \left(0, 2, \frac{\pi}{2}\right) \cdot \frac{2}{\sqrt{5}}\left(-1, 0, \frac{1}{2}\right) = \frac{\pi}{2\sqrt{5}}.$$

Así pues, la razón de cambio de f respecto a la distancia a lo largo de $\mathscr{C}$ en $\left(0, 1, \frac{\pi}{4}\right)$ es $\frac{\pi}{2\sqrt{5}} = \frac{\sqrt{5}}{10}\pi$.

La diferencial es a menudo una aproximación práctica y muy exacta de un incremento (problemas 8-10). Haciendo $\Delta f(\mathbf{x}; \mathbf{h}) = f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})$, tenemos, si f es diferenciable en $\mathbf{x}$,

$$\Delta f(\mathbf{x}; \mathbf{h}) = df(\mathbf{x}; \mathbf{h}) + \mathbf{h} \cdot \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) \text{ donde } \lim_{\mathbf{h} \to \mathbf{0}} \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) = \mathbf{0}.$$

Así pues, si la longitud de $\mathbf{h}$ es "pequeña", $df(\mathbf{x}; \mathbf{h})$ es aproximadamente igual al incremento $\Delta f(\mathbf{x}; \mathbf{h})$. Más adelante, en la sección 11, podremos establecer cotas para el error que se comete en este tipo de aproximaciones.

Problemas

1. Determínese el gradiente de f cuando

$a)$ $f = I_1{}^2 + I_2{}^2 + I_3{}^2$ $\qquad$ $b)$ $f(x, y, z) = \dfrac{x-y}{x+z}$

$c)$ $f = I_1{}^2 I_2 I_3{}^2$ $\qquad$ $d)$ $f(x, y) = e^{xy^2}$

$e)$ $f(x, y) = e^x \cos(x+y)$ $\qquad$ $f)$ $f(x, y, z) = \ln(x^2+y^2+z^2)$.

2. Si $w = f(\mathbf{x})$, encuéntrese $dw = df(\mathbf{x}; d\mathbf{x})$ cuando

$a)$ $f(\mathbf{x}) = f(x, y) = x^2 y^3 + 3y$

$b)$ $f(\mathbf{x}) = f(x, y) = \arctan \dfrac{y}{x}$

$c)$ $f(\mathbf{x}) = f(x, y, z) = \sqrt{x^2+y^2+z^2}$

$d)$ $f(\mathbf{x}) = f(x, y, z) = z\cos^2(x+y)$.

3. Determínese $D_{\mathbf{u}} f$ cuando

$a)$ $f(x, y) = \dfrac{x^2+y}{x-y}$, $\qquad$ $\mathbf{u} = \dfrac{1}{\sqrt{10}}(-3, 1)$

$b)$ $f = I_2{}^2 - I_1{}^3 I_3$, $\qquad$ $\mathbf{u} = \dfrac{1}{3\sqrt{5}}(5, -2, 4)$

$c)$ $f(x, y, z) = \dfrac{xy}{z}$, $\qquad$ $\mathbf{u} = \dfrac{1}{\sqrt{3}}(1, 1, 1)$

$d)$ $f(x, y, z, t) = \dfrac{x+y+z-t}{x+z}$, $\qquad$ $\mathbf{u} = \dfrac{1}{\sqrt{10}}(2, -1, 2, 1)$.

4. Encuéntrese la razón de cambio máxima de las siguientes funciones en los puntos que en cada caso se indican:

$a)$ $f = I_1 I_2{}^2 + I_1{}^2 I_3$; $(3, 1, 2)$

$b)$ $f = \dfrac{I_1 I_3}{I_1 + I_2}$; $(1, -2, 0)$

$c)$ $f(x, y, z) = \sqrt{x^2+y^2+z^2}$; $(3, 4, -3)$

$d)$ $f(x, y, z, t) = xz + y^2 t$; $(1, 0, -3, 2)$.

5. Supongamos que $w = f(x, y, z)$ y $x = g_1(t)$, $y = g_2(t)$, $z = g_3(t)$.

Entonces $w = f(g_1(t), g_2(t), g_3(t)) = F(t)$. Demuéstrese que bajo apropiadas condiciones

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt}$$

y exprésense $\frac{dw}{dt}$ y $\frac{\partial w}{\partial x}$ en notación funcional.

6. Si $w = xy^2 + z^2$ y $x = \cos t$, $y = e^t$, $z = t^2$, determínese $\frac{dw}{dt}$ usando la fórmula del problema 5 y también expresando w en términos de t y diferenciando a continuación.

7. Si $f(x, y) = \dfrac{x^2 y}{x^2 + y^2}$, determínense $df((1,2);(\frac{1}{2},\frac{1}{4}))$ y $\Delta f((1,2);(\frac{1}{2},\frac{1}{4}))$.

8. Un tanque cilíndrico de hierro tiene 6 pies de altura y un diámetro exterior de 2 pies. Si la tapa y la parte inferior del tanque tienen $\frac{1}{2}$ pulgada de grosor y las paredes $\frac{1}{4}$ de pulgada, úsese la diferencial para calcular en forma aproximada el peso del tanque si el peso del hierro es de 450 libras por pie cúbico.

9. El resultado de medir una habitación nos indica que sus dimensiones son $12 \times 16 \times 8$ pies. Sabemos que el error cometido en cada una de ellas es menor de 1 pulgada. Calcúlese aproximadamente el máximo error posible que ese comete al aceptar como volumen de la habitación el anterior producto.

10. El voltaje E en una resistencia R y la corriente I que pasa por R están relacionados de acuerdo con la ley de Ohm: $I = \dfrac{E}{R}$. Si se sabe que la resistencia R es de 100 000 ohms $\pm 2\%$ y E es igual a 150 volts con una variación posible de 5 volts, calcúlese en forma aproximada cuál es el máximo error posible que se comete al calcular I tomando R igual a 100 000 ohms y E igual a 150 volts.

10. DERIVADAS PARCIALES DE ORDEN SUPERIOR

Como la derivada parcial $D_k f$ de una función f de R^n en R es asimismo una función de R^n en R, podemos tomar derivadas parciales de $D_k f$. La derivada parcial de $D_k f$ con respecto a la coordenada j-ésima es $D_j(D_k f)$. Esta se denotará por $D_{j,k} f$. La función $D_{j,k} f$ se llama derivada parcial segunda de f. Es claro que podemos seguir tomando derivadas parciales y obtener derivadas parciales de f de órdenes más altos.

En nuestra notación para derivadas parciales, si hacemos

$$w = f(\mathbf{x}) = f(x_1, \ldots, x_n),$$

entonces

$$D_{j,k}f(\mathbf{x}) = \frac{\partial}{\partial x_j}\left(\frac{\partial w}{\partial x_k}\right) = \frac{\partial^2 w}{\partial x_j \partial x_k}.$$

10.1 Ejemplo. Si f está definida por $f(x, y) = x^2+3xy$, encuéntrense las derivadas parciales de segundo orden de f.

SOLUCIÓN. Sea $z = f(x, y)$. Entonces

$$\frac{\partial z}{\partial x} = D_1 f(x, y) = 2x+3y, \qquad \frac{\partial z}{\partial y} = D_2 f(x, y) = 3x$$

$$\frac{\partial^2 z}{\partial x^2} = D_{1,1} f(x, y) = 2, \qquad \frac{\partial^2 z}{\partial y^2} = D_{2,2} f(x, y) = 0$$

$$\frac{\partial^2 z}{\partial y\, \partial x} = D_{2,1} f(x, y) = 3, \qquad \frac{\partial^2 z}{\partial x\, \partial y} = D_{1,2} f(x, y) = 3.$$

Nótese que en este ejemplo $D_{2,1} f = D_{1,2} f$. Demostraremos que esta igualdad de las derivadas parciales mixtas se verifica en una gran clase de funciones.

Antes de dar un teorema sobre la igualdad de las derivadas parciales mixtas de una función de R^2 en R, probaremos un lema que usamos en la demostración de tal teorema.

10.2 Lema. *Si $D_{2,1} f$ existe en el interior y en la frontera del rectángulo $\mathscr{R}$ con vértices (x_0, y_0), (x_0+h, y_0), (x_0+h, y_0+k), (x_0, y_0+k), entonces*

$$[f(x_0+h, y_0+k)-f(x_0+h, y_0)] \\ -[f(x_0, y_0+k)-f(x_0, y_0)] = hkD_{2,1} f(x, y)$$

para algún punto (x, y) en $\mathscr{R}$.

PRUEBA. Sea

$$g(x) = f(x, y_0+k)-f(x, y_0).$$

Entonces

$$g'(x) = D_1 f(x, y_0+k)-D_1 f(x, y_0).$$

Como $D_1 f$ existe sobre $\mathscr{R}$, g' existe sobre $[x_0, x_0+h]$ ($[x_0+h, x_0]$ si $h < 0$). Usando el teorema del valor medio tenemos

$$\begin{aligned}[f(x_0+h, y_0+k)-f(x_0+h, y_0)]-[f(x_0, y_0+k)-f(x_0, y_0)] \\ &= g(x_0+h)-g(x_0) \\ &= hg'(x) \\ &= h[D_1 f(x, y_0+k)-D_1 f(x, y_0)]\end{aligned}$$

para algún x entre x_0 y x_0+h. Entonces, aplicando el teorema del valor medio a la función $D_1 f$,

$$\begin{aligned}[f(x_0+h, y_0+k)-f(x_0+h, y_0)]-[f(x_0, y_0+k)-f(x_0, y_0)] \\ &= h[D_1 f(x, y_0+k)-D_1 f(x, y_0)] \\ &= hkD_{2,1} f(x, y)\end{aligned}$$

para algún y entre y_0 y y_0+k. Lo que completa la prueba del lema.

Enunciamos y probamos ahora el teorema sobre la igualdad de las derivadas parciales mixtas de una función de R^2 en R.

10.3 Teorema. *Si $D_{2,1} f$ y $D_2 f$ existen en una vecindad $\mathscr{S}(\mathbf{x}_0; r)$ de $\mathbf{x}_0 = (x_0, y_0)$ y $D_{2,1} f$ es continua en $\mathbf{x}_0$, entonces $D_{1,2} f$ existe en $\mathbf{x}_0$ y $D_{1,2} f(\mathbf{x}_0) = D_{2,1} f(x_0)$.*

Prueba. Por definición

$$D_{1,2} f(\mathbf{x}_0) = \lim_{h\to 0} \frac{D_2 f(x_0+h, y_0)-D_2 f(x_0, y_0)}{h}$$

$$= \lim_{h\to 0} \lim_{k\to 0} \frac{[f(x_0+h, y_0+k)-f(x_0+h, y_0)]-[f(x_0, y_0+k)-f(x_0, y_0)]}{hk}$$

si este límite iterado existe. Como $D_2 f$ existe en $\mathscr{S}(\mathbf{x}_0; r)$, cuando $0 < |h| < r$

$$\lim_{k\to 0} \frac{[f(x_0+h, y_0+k)-f(x_0+h, y_0)]-[f(x_0, y_0+k)-f(x_0, y_0)]}{hk}$$

existe. Llamemos a este límite $g(h)$. Ahora demostraremos que $\lim_{h\to 0} g(h)$ existe y es igual a $D_{2,1} f(\mathbf{x}_0)$. Tomemos $\varepsilon > 0$. Como $D_{2,1} f$ es continua en $\mathbf{x}_0$, existe un $\delta > 0$ tal que $\delta \leqslant r$ y

$$|D_{2,1} f(\mathbf{x})-D_{2,1} f(\mathbf{x}_0)| < \tfrac{\varepsilon}{2} \text{ siempre que } |\mathbf{x}-\mathbf{x}_0| < \delta.$$

Tomemos h tal que $0 < |h| < \delta$. Entonces tomemos $k \neq 0$ tal que $\sqrt{h^2+k^2}$

y

$$\left| g(h) - \frac{[f(x_0+h, y_0+k)-f(x_0+h, y_0)]-[f(x_0, y_0+k)-f(x_0, y_0)]}{hk} \right|$$

$$< \frac{\varepsilon}{2}.$$

Para h y k tales, el lema 10.2 implica que existe un punto $\mathbf{x}$ en el rectángulo con vértices (x_0, y_0), (x_0+h, y_0), (x_0+h, y_0+k), (x_0, y_0+k) tal que

$$\frac{[f(x_0+h, y_0+k)-f(x_0+h, y_0)]-[f(x_0, y_0+k)-f(x_0, y_0)]}{hk} = D_{2,1} f(\mathbf{x}).$$

Nótese que este punto $\mathbf{x}$ está en $\mathscr{S}(\mathbf{x}_0; \delta)$. Tenemos

$$|g(h)-D_{2,1} f(\mathbf{x}_0)|$$

$$\leqslant \left| g(h) - \frac{[f(x_0+h, y_0+k)-f(x_0+h, y_0)]-[f(x_0, y_0+k)-f(x_0, y_0)]}{hk} \right|$$

$$+ |D_{2,1} f(\mathbf{x})-D_{2,1} f(\mathbf{x}_0)|$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Lo que nos muestra que $\lim_{h \to 0} g(h) = D_{2,1} f(\mathbf{x}_0)$ y completa la prueba.

El teorema 10.3 puede aplicarse a derivadas parciales mixtas de orden mayor que dos. Pasamos ahora, antes de dar algunos ejemplos, a definir un nuevo concepto.

10.4 Definición. *Se dice que una función f pertenece a la clase C^n sobre un conjunto abierto $\mathscr{E}$, lo que escribiremos: $f \in C^n$ sobre $\mathscr{E}$, si todas las derivadas parciales de orden n de f son continuas sobre $\mathscr{E}$.*

Supongamos que $f \in C^3$ sobre algún conjunto abierto $\mathscr{E}$ de R^2. Demostraremos que $D_{2,2,1} f = D_{1,2,2} f$ sobre $\mathscr{E}$. Como todas las derivadas parciales de tercer orden de f son continuas sobre $\mathscr{E}$, todas las derivadas parciales de orden más bajo de f existen y son continuas sobre $\mathscr{E}$ (teoremas 8.6, pág. 205, y 6.5, pág. 190). Por tanto, aplicando el teorema 10.3 a f, tenemos $D_{2,1} f = D_{1,2} f$ sobre $\mathscr{E}$ y por tanto

$$D_{2,2,1} f = D_{2,1,2} f \quad \text{sobre } \mathscr{E}.$$

Aplicando el teorema 10.3 a $D_2 f$ tenemos

$$D_{2,1,2} f = D_{2,1}(D_2 f) = D_{1,2}(D_2 f) = D_{1,2,2} f \quad \text{sobre } \mathscr{E}.$$

Por tanto, $D_{2,2,1} f = D_{1,2,2} f$ sobre $\mathscr{E}$.

Si hacemos $z = f(x, y)$, entonces podemos escribir esto en la forma

$$\frac{\partial^3 z}{\partial y^2 \partial x} = \frac{\partial^3 z}{\partial y \partial y \partial x} = \frac{\partial^3 z}{\partial x \partial y \partial y} = \frac{\partial^3 z}{\partial x \partial y^2}.$$

El teorema 10.3 puede también aplicarse a derivadas parciales de funciones definidas en R^3 o, en general, en R^n. Por ejemplo, si $f \in C^2$ sobre un conjunto abierto $\mathscr{E}$ en R^3, entonces podemos demostrar que $D_{1,2} f = D_{2,1} f$ sobre $\mathscr{E}$. Para cualquier punto (x, y, z_0) en $\mathscr{E}$, sea $g(x, y) = f(x, y, z_0)$. Entonces

$$D_{1,2} f(x, y, z_0) = D_{1,2} g(x, y) = D_{2,1} g(x, y) = D_{2,1} f(x, y, z_0).$$

Por tanto, $D_{1,2} f = D_{2,1} f$ sobre $\mathscr{E}$.

En general, si $f \in C^n$ sobre un conjunto abierto $\mathscr{E}$, podemos decir que cualesquier dos derivadas parciales de orden n o menor, con los mismos subíndices, son iguales en $\mathscr{E}$ aunque el orden en que los subíndices aparecen sea diferente.

El hecho de que las derivadas parciales mixtas de una función no son siempre iguales se demuestran en el siguiente ejemplo.

10.5 Ejemplo. Encuéntrense $D_{2,1} f(0, 0)$ y $D_{1,2} f(0, 0)$ cuando

$$f(x, y) = xy \frac{x^2 - y^2}{x^2 + y^2}, \quad (x, y) \neq (0, 0)$$

$$f(0, 0) = 0.$$

SOLUCIÓN.

$$D_1 f(x, y) = xy \frac{2x(x^2+y^2) - 2x(x^2-y^2)}{(x^2+y^2)^2} + y \frac{x^2-y^2}{x^2+y^2}$$

$$= \frac{4x^2 y^3}{(x^2+y^2)^2} + y \frac{x^2-y^2}{x^2+y^2}, \quad (x, y) \neq (0, 0)$$

$$D_1 f(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \to 0} 0 = 0$$

$$D_2 f(x, y) = xy \frac{-2y(x^2+y^2) - 2y(x^2-y^2)}{(x^2+y^2)^2} + x \frac{x^2-y^2}{x^2+y^2}$$

$$= \frac{-4x^3 y^2}{(x^2+y^2)^2} + x \frac{x^2-y^2}{x^2+y^2}, \quad (x, y) \neq (0, 0)$$

$$D_2 f(0, 0) = \lim_{h \to 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \to 0} 0 = 0$$

$$D_{1,2} f(0, 0) = \lim_{h \to 0} \frac{D_2 f(h, 0) - D_2 f(0, 0)}{h} = \lim_{h \to 0} \frac{h}{h} = 1$$

$$D_{2,1} f(0, 0) = \lim_{h \to 0} \frac{D_1 f(0, h) - D_1 f(0, 0)}{h} = \lim_{h \to 0} -\frac{h}{h} = -1 .$$

Así pues, en este ejemplo $D_{1,2} f(0, 0) \neq D_{2,1} f(0, 0)$.

Problemas

1. Determínense todas las derivadas parciales de segundo orden de f cuando

a) $f(x, y) = xy^2 + e^{xy}$ *b*) $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$

c) $f(x, y) = \tan \dfrac{y}{x}$ *d*) $f = \exp \circ (I_1{}^2 I_2)$.

2. Encuéntrense los valores de las derivadas parciales de segundo orden de f en los puntos que se indican:

a) $f(x, y) = x^2 y^3 + x^3 y$; $(-3, 2)$

b) $f = \ln \circ (I_1{}^2 + I_2{}^2 + I_3{}^2)$; $(-2, \sqrt{3}, 2)$

c) $f(x, y, z) = \dfrac{x^2 y}{y - z}$; $(\frac{1}{2}, 3, -1)$.

3. Determínense todas las derivadas parciales de tercer orden de f cuando

a) $f(x, y) = \cos xy$ *b*) $f(x, y, z) = x^3 z + yz$.

4. Si $f(x, y) = x^2 y \operatorname{sen} \dfrac{1}{x}$ para $x \neq 0$ y $f(0, y) = 0$, determínense $D_{2,1} f(0, 0)$ y $D_{1,2} f(0, 0)$. ¿Es continua $D_{2,1} f$ en $(0, 0)$?

11. EL TEOREMA DE TAYLOR

En esta sección extendemos el teorema de Taylor a funciones reales de un vector. Se obtiene fácilmente esta extensión partiendo del teorema de Taylor para funciones reales de una variable real.

El teorema de Taylor para funciones de R en R dice: *si f tiene derivadas continuas hasta la de orden $n+1$ sobre el intervalo $\mathscr{J}$ y $x_0 \in \mathscr{J}$, entonces para cualquier x en $\mathscr{J}$ distinta de x_0*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_n$$

donde

$$R_n = \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$$

para algún c entre x_0 y x. El símbolo $f^{(0)}$ que anteriormente aparece denota a f, es decir, $f^{(0)} = f$. Este teorema da una expresión (la de R_n) del error cometido al aproximarnos a la función f por el polinomio

$$\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(I-x_0)^k$$

en el intervalo $\mathscr{J}$.

Sea ahora f una función de R^2 en R y sean $\mathbf{x}_0$ y $\mathbf{x}$ dos puntos en el dominio de f. Si nos limitamos a la consideración de los valores de la función sobre el segmento rectilíneo $[\mathbf{x}_0, \mathbf{x}]$, entonces podemos considerar f como una función de una sola variable real; nos basta definir g por la regla

$$g(t) = f(\mathbf{x}_0 + t(\mathbf{x}-\mathbf{x}_0)) \quad \text{donde} \quad t \in [0, 1].$$

Aplicando luego el teorema de Taylor a g sobre el intervalo $[0, 1]$ (si es que puede hacerse), obtenemos

11.1
$$g(1) = \sum_{k=0}^{n} \frac{g^{(k)}(0)}{k!} + R_n$$

donde

$$R_n = \frac{g^{(n+1)}(\theta)}{(n+1)!} \quad \text{para algún } \theta \in \langle 0, 1 \rangle.$$

Esto dará lugar a la fórmula de Taylor para f cuando reemplacemos g y sus derivadas por las expresiones apropiadas en f y sus derivadas parciales.

Investigamos ahora qué condiciones son las que debe exigirse que cumpla f para que pueda obtenerse 11.1. La expresión 11.1 es válida si g tiene derivadas continuas hasta de orden $n+1$ sobre $[0, 1]$. Debemos, pues, encontrar la relación existente entre las derivadas de g y las derivadas parciales de f. Sea $\mathbf{x}-\mathbf{x}_0 = \mathbf{h} = (h_1, h_2)$. Entonces, $g(t) = f(\mathbf{x}_0 + t\mathbf{h})$. Usando la regla de la cadena (6.9, pág. 193), obtenemos

$$g'(t) = \mathbf{D}f(\mathbf{x}_0 + t\mathbf{h}) \cdot \mathbf{h} = h_1 D_1 f(\mathbf{x}_0 + t\mathbf{h}) + h_2 D_2 f(\mathbf{x}_0 + t\mathbf{h})$$

y

$$\begin{aligned} g''(t) &= h_1 D_1 [h_1 D_1 f(\mathbf{x}_0 + t\mathbf{h}) + h_2 D_2 f(\mathbf{x}_0 + t\mathbf{h})] \\ &\qquad + h_2 D_2 [h_1 D_1 f(\mathbf{x}_0 + t\mathbf{h}) + h_2 D_2 f(\mathbf{x}_0 + t\mathbf{h})] \\ &= h_1^2 D_{1,1} f(\mathbf{x}_0 + t\mathbf{h}) + h_1 h_2 D_{1,2} f(\mathbf{x}_0 + t\mathbf{h}) \\ &\qquad + h_2 h_1 D_{2,1} f(\mathbf{x}_0 + t\mathbf{h}) + h_2^2 D_{2,2} f(\mathbf{x}_0 + t\mathbf{h}). \end{aligned}$$

Así pues, si $f \in C^2$ sobre un conjunto abierto $\mathscr{E}$ que contiene el segmento $[\mathbf{x}_0, \mathbf{x}_0 + \mathbf{h}]$, entonces g tendrá derivadas continuas hasta de orden 2 sobre $[0, 1]$. Además, si $f \in C^2$ sobre $\mathscr{E}$, entonces $D_{1,2} f = D_{2,1} f$ sobre $[\mathbf{x}_0, \mathbf{x}_0 + \mathbf{h}]$ y, por tanto

$$g''(t) = h_1^2 D_{1,1} f(\mathbf{x}_0 + t\mathbf{h}) + 2h_1 h_2 D_{1,2} f(\mathbf{x}_0 + t\mathbf{h}) + h_2^2 D_{2,2} f(\mathbf{x}_0 + t\mathbf{h}).$$

En general, si suponemos que $f \in C^{n+1}$ sobre un conjunto abierto que contiene el segmento $[\mathbf{x}_0, \mathbf{x}_0 + \mathbf{h}]$, entonces, para $k = 1, \ldots, n+1$, $g^{(k)}$ es continua sobre $[0, 1]$ y

11.2
$$g^{(k)}(t) = \sum_{i=0}^{k} \binom{k}{i} h_1^{k-i} h_2^{\,i} D_1^{k-i} D_2^{\,i} f(\mathbf{x}_0 + t\mathbf{h})$$

donde $\binom{k}{i}$ es el coeficiente del binomio, es decir, $\binom{k}{i} = \dfrac{k!}{i!(k-i)!}$. El segundo miembro de 11.2 es análogo a una expresión en el teorema del binomio y la prueba de 11.2 por inducción matemática es paralela a la del teorema del binomio. Con el teorema del binomio en mente, es natural escribir 11.2 de la siguiente forma

11.3
$$g^{(k)}(t) = [h_1 D_1 + h_2 D_2]^k f(\mathbf{x}_0 + t\mathbf{h}), \quad k = 1, \ldots, n+1.$$

Si hacemos $[h_1 D_1 + h_2 D_2]^0 f(\mathbf{x}_0 + t\mathbf{h}) = f(\mathbf{x}_0 + t\mathbf{h})$, entonces 11.3 se verifica para $k = 0, \ldots, n+1$.

Damos ahora el teorema de Taylor para funciones de R^2 en R.

11.4 Teorema. (Teorema de Taylor.) *Si f es una función de R^2 en R que pertenece a la clase C^{n+1} sobre un conjunto abierto $\mathscr{E}$ y si $\mathbf{x}_0 \in \mathscr{E}$, entonces para cualquier $\mathbf{x} \in \mathscr{E}$, distinto de $\mathbf{x}_0$, tal que $[\mathbf{x}_0, \mathbf{x}] \subset \mathscr{E}$*

11.5
$$f(\mathbf{x}) = \sum_{k=0}^{n} \frac{1}{k!} [(x - x_0) D_1 + (y - y_0) D_2]^k f(\mathbf{x}_0) + R_n$$

donde

$$R_n = \frac{1}{(n+1)!} [(x - x_0) D_1 + (y - y_0) D_2]^{n+1} f(\mathbf{c}) \quad \textit{para algún } \mathbf{c} \in \langle \mathbf{x}_0, \mathbf{x} \rangle.$$

PRUEBA. Sea $g(t) = f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))$, $t \in [0, 1]$. Entonces, para $k = 1, \ldots, n+1$

$$g^{(k)}(t) = [(x - x_0) D_1 + (y - y_0) D_2]^k f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)).$$

Así pues, g tiene derivadas continuas hasta de orden $n+1$ sobre $[0, 1]$ y el teorema de Taylor aplicado a g sobre $[0, 1]$ nos da

$$g(1) = \sum_{k=0}^{n} \frac{g^{(k)}(0)}{k!} + R_n$$

donde $R_n = \dfrac{g^{(n+1)}(\theta)}{(n+1)!}$ para cierto $\theta \in \langle 0, 1 \rangle$.

Tenemos, por tanto,

$$f(\mathbf{x}) = \sum_{k=0}^{n} \frac{1}{k!} [(x-x_0)D_1 + (y-y_0)D_2]^k f(\mathbf{x}_0) + R_n$$

donde

$$R_n = \frac{1}{(n+1)!} [(x-x_0)D_1 + (y-y_0)D_2]^{n+1} f(\mathbf{c}) \quad \text{para algún } \mathbf{c} \in \langle \mathbf{x}_0, \mathbf{x} \rangle.$$

Lo que completa la prueba.

La fórmula 11.5 en el teorema de Taylor se llama *fórmula de Taylor* en $\mathbf{x}_0$ y al término R_n se le llama el *residuo*. Si hacemos $n = 1$ en la fórmula de Taylor obtenemos una fórmula para la aproximación por diferenciales:

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{x}_0) &= (x-x_0)D_1 f(\mathbf{x}_0) + (y-y_0)D_2 f(\mathbf{x}_0) + R_1 \\ &= (\mathbf{x}-\mathbf{x}_0) \cdot \nabla f(\mathbf{x}_0) + R_1 \\ &= df(\mathbf{x}_0; \mathbf{x}-\mathbf{x}_0) + R_1 \end{aligned}$$

donde

$$R_1 = \tfrac{1}{2}[(x-x_0)^2 D_{1,1} f(\mathbf{c}) + 2(x-x_0)(y-y_0) D_{1,2} f(\mathbf{c}) + (y-y_0)^2 D_{2,2} f(\mathbf{c})]$$

para cierto $\mathbf{c} \in \langle \mathbf{x}_0, \mathbf{x} \rangle$.

Nota. Obsérvese que si introducimos la expresión simbólica

$$(x-x_0)D_1 + (y-y_0)D_2 = (\mathbf{x}-\mathbf{x}_0) \cdot \mathbf{D}$$

entonces el teorema de Taylor toma la forma

$$f(\mathbf{x}) = \sum_{k=0}^{n} \frac{1}{k!} [(\mathbf{x}-\mathbf{x}_0) \cdot \mathbf{D}]^k f(\mathbf{x}_0) + R_n$$

donde

$$R_n = \frac{1}{(n+1)!} [(\mathbf{x}-\mathbf{x}_0) \cdot \mathbf{D}]^{n+1} f(\mathbf{c}) \quad \text{para algún } \mathbf{c} \in \langle \mathbf{x}_0, \mathbf{x} \rangle.$$

El teorema de Taylor para funciones de R^n en R puede escribirse en esta misma forma pero la prueba sería más complicada.

11.6 Ejemplo. Desarróllese la fórmula de Taylor en $\mathbf{x}_0 = (0, 0)$ y con $n = 3$ para la función f definida por $f(x, y) = e^x \cos y$.

SOLUCIÓN. Las derivadas parciales de f hasta el orden cuatro son:

$$D_1 f(x, y) = e^x \cos y, \; D_2 f(x, y) = -e^x \operatorname{sen} y$$
$$D_{1,1} f(x, y) = e^x \cos y, D_{1,2} f(x, y) = -e^x \operatorname{sen} y, D_{2,2} f(x, y) = -e^x \cos y$$
$$D_{1,1,1} f(x, y) = e^x \cos y, \; D_{1,1,2} f(x, y) = -e^x \operatorname{sen} y,$$
$$D_{1,2,2} f(x, y) = -e^x \cos y, \; D_{2,2,2} f(x, y) = e^x \operatorname{sen} y,$$
$$D_{1,1,1,1} f(x, y) = e^x \cos y, \; D_{1,1,1,2} f(x, y) = -e^x \operatorname{sen} y,$$
$$D_{1,1,2,2} f(x, y) = -e^x \cos y, \; D_{1,2,2,2} f(x, y) = e^x \operatorname{sen} y,$$
$$D_{2,2,2,2} f(x, y) = e^x \cos y.$$

Así pues, $f \in C^4$ en R^2 y, por tanto, para cualquier punto $(x, y) \in \mathrm{R}^2$,

$$f(x, y) = [xD_1 + yD_2]^0 f(0,0) + [xD_1 + yD_2] f(0,0) + \frac{1}{2!}[xD_1 + yD_2]^2 f(0, 0)$$
$$+ \frac{1}{3!}[xD_1 + yD_2]^3 f(0, 0) + R_3$$
$$= f(0, 0) + xD_1 f(0, 0) + yD_2 f(0, 0)$$
$$+ \tfrac{1}{2}[x^2 D_{1,1} f(0, 0) + 2xyD_{1,2} f(0, 0) + y^2 D_{2,2} f(0, 0)]$$
$$+ \tfrac{1}{6}[x^3 D_{1,1,1} f(0,0) + 3x^2 y D_{1,1,2} f(0, 0) + 3xy^2 D_{1,2,2} f(0, 0)$$
$$+ y^3 D_{2,2,2} f(0, 0)] + R_3$$
$$= 1 + x + \tfrac{1}{2}(x^2 - y^2) + \tfrac{1}{6}(x^3 - 3xy^2) + R_3$$

donde $R_3 = \dfrac{1}{4!}[xD_1 + yD_2]^4 f(c_1, c_2)$

$$= \frac{e^{c_1}}{24}(x^4 \cos c_2 - 4x^3 y \operatorname{sen} c_2 - 6x^2 y^2 \cos c_2 + 4xy^3 \operatorname{sen} c_2 + y^4 \cos c_2)$$

para un cierto $(c_1, c_2) \in \langle (0, 0), (x, y) \rangle$.

Problemas

1. Escríbase la fórmula de Taylor en los siguientes casos, especificando para qué puntos (x, y) tiene validez.

a) $f(x, y) = x^2 - 3xy + 4y^2$, $\mathbf{x}_0 = (0, 0)$, $n = 2$
b) $f(x, y) = x^2 - 3xy + 4y^2$, $\mathbf{x}_0 = (2, -3)$, $n = 2$

c) $f(x, y) = 2x^4 - 5y^3 + 2xy^2$, $\mathbf{x}_0 = (0, 0)$, $n = 2$
d) $f(x, y) = 2x^4 - 5y^3 + 2xy^2$, $\mathbf{x}_0 = (0, 0)$, $n = 4$
e) $f(x, y) = y^2 \ln x$, $\mathbf{x}_0 = (1, 0)$, $n = 4$
f) $f(x, y) = \cos x \cos y$, $\mathbf{x}_0 = (0, 0)$, $n = 6$
g) $f(x, y) = \cos xy$, $\mathbf{x}_0 = (0, 0)$, $n = 6$
h) $f(x, y) = e^{xy} \operatorname{sen} y$, $\mathbf{x}_0 = (0, 0)$, $n = 3$
i) $f(x, y) = \sqrt{x^3 y}$, $\mathbf{x}_0 = (1, 3)$, $n = 3$.

2. Pruébese 11.2 por inducción matemática.

3. Determínense los siguientes límites:

a) $\lim\limits_{(x,y)\to(0,0)} \dfrac{\operatorname{sen} xy}{y}$

b) $\lim\limits_{(x,y)\to(0,0)} \dfrac{\tan xy}{x}$

c) $\lim\limits_{(x,y)\to(0,0)} \dfrac{e^{xy} - 1}{x}$.

12. PLANO TANGENTE A UNA SUPERFICIE

En este capítulo hemos estado usando el término superficie sin haber expresado en forma precisa qué es lo que entendemos por tal cosa. Llamamos superficie a la gráfica de una función f de R^2 en R. Este es un conjunto de la forma

12.1 $$\{(x, y, z) \mid z = f(x, y)\}.$$

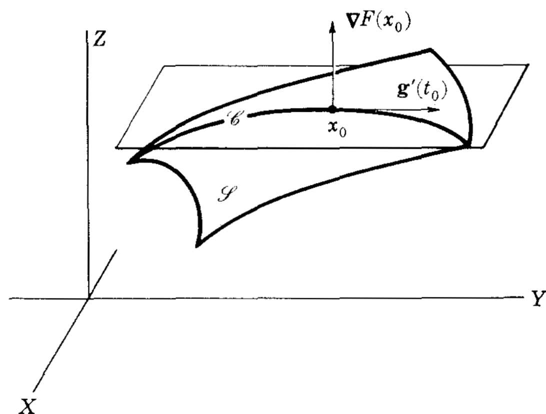


FIGURA 16

Llamamos también "superficies de nivel", al discutir la representación de una función F de R^3 en R, a conjuntos de la forma

12.2 $$\{(x, y, z) \mid F(x, y, z) = c\}.$$

Los conjuntos de la forma 12.1 o 12.2 que encontramos tienen una propiedad común que podemos describir diciendo que un punto puede moverse en el conjunto con dos grados de libertad. Dejamos por el momento la noción de superficie en este vago estado. Aunque seremos un poco más precisos sobre este punto en el próximo capítulo, una descripción completa del concepto es bastante complicada y se encontraría aquí fuera de lugar.

Es claro que cualquier conjunto de la forma 12.1 puede escribirse en la forma 12.2; simplemente, haciendo $F(x, y, z) = f(x, y) - z$. Bajo ciertas circunstancias (que discutiremos en la próxima sección) un conjunto de la forma 12.2 puede escribirse en la forma 12.1 .

Sea F una función de R^3 en R que es diferenciable sobre un conjunto abierto $\mathscr{E}$ y sea $\mathscr{S}$ la superficie

$$\mathscr{S} = \{(x, y, z) \in \mathscr{E} \mid F(x, y, z) = c\}.$$

Sea $\mathbf{x}_0 = (x_0, y_0, z_0)$ un punto sobre $\mathscr{S}$ y sea

$$\mathbf{x} = \mathbf{g}(t), \; t \in \langle a, b \rangle$$

la ecuación de una curva $\mathscr{C}$ sobre $\mathscr{S}$ que pasa por $\mathbf{x}_0$ (figura 16). Como $\mathbf{x}_0$ está sobre $\mathscr{C}$, $\mathbf{x}_0 = \mathbf{g}(t_0)$ para un cierto $t_0 \in \langle a, b \rangle$. Y como $\mathscr{C}$ se encuentra sobre $\mathscr{S}$,

$$F(\mathbf{g}(t)) = c \quad \text{para todo} \quad t \in \langle a, b \rangle.$$

Si suponemos que $\mathbf{g}$ es diferenciable sobre $\langle a, b \rangle$, entonces, de acuerdo con el teorema 6.9 (pág. 193), $F \circ \mathbf{g}$ es diferenciable sobre $\langle a, b \rangle$ y

$$\nabla F(\mathbf{g}(t)) \cdot \mathbf{g}'(t) = 0 \quad \text{para todo} \quad t \in \langle a, b \rangle.$$

En particular, cuando $t = t_0$

12.3 $$\nabla F(\mathbf{g}(t_0)) \cdot \mathbf{g}'(t_0) = 0.$$

Según la ecuación 12.3 vemos que *en* $\mathbf{x}_0$ *el gradiente de* F *es ortogonal al vector tangente a cualquier curva* $\mathbf{x} = \mathbf{g}(t)$ *que se encuentre sobre la superficie* $\mathscr{S}$ *y pase por* $\mathbf{x}_0$. *Así, pues, si* $\nabla F(\mathbf{x}_0) \neq \mathbf{0}$, *las tangentes de todas las curvas sobre* $\mathscr{S}$ *en el punto* $\mathbf{x}_0$ *se encuentran sobre un mismo plano.*

Si $\nabla F(\mathbf{x}_0) \neq \mathbf{0}$, definimos como *plano tangente* a la superficie

$$\mathscr{S} = \{(x, y, z) \mid F(x, y, z) = c\}$$

en el punto $\mathbf{x}_0$ al plano que pasa por $\mathbf{x}_0$ y tiene como normal a $\nabla F(\mathbf{x}_0)$. Así pues, el plano tangente a $\mathscr{S}$ en $\mathbf{x}_0$ tiene la ecuación

12.4 $$(\mathbf{x} - \mathbf{x}_0) \cdot \nabla F(\mathbf{x}_0) = 0$$

o, en otra notación,

$$(x-x_0)\frac{\partial F}{\partial x}+(y-y_0)\frac{\partial F}{\partial y}+(z-z_0)\frac{\partial F}{\partial z}=0$$

donde las derivadas parciales son evaluadas en el punto (x_0, y_0, z_0). Si $\nabla F(\mathbf{x}_0) = \mathbf{0}$, entonces $\mathscr{S}$ no tiene plano tangente en $\mathbf{x}_0$.

12.5 Ejemplo. Determínese una ecuación del plano tangente a la superficie $\mathscr{S}$ de la ecuación

$$\frac{x^2}{4}+\frac{y^2}{3}-z^2=1$$

en el punto $(2, 3, -\sqrt{3})$.

SOLUCIÓN. La superficie $\mathscr{S}$ se llama hiperboloide de una hoja (figura 17).

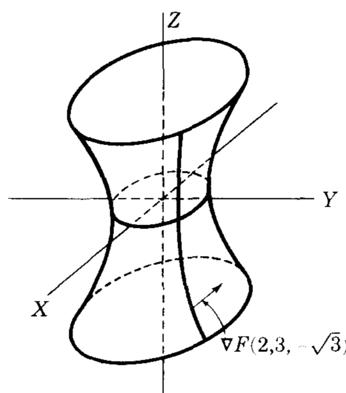


FIGURA 17

Haciendo $F(x, y, z) = \frac{x^2}{4}+\frac{x^2}{3}-z^2$, tenemos

$$\nabla F(x, y, z) = (\tfrac{1}{2}x, \tfrac{2}{3}y, -2z)$$

y

$$\nabla F(2, 3, -\sqrt{3}) = (1, 2, 2\sqrt{3}).$$

Por tanto, una ecuación del plano tangente es

$$(1, 2, 2\sqrt{3})\cdot[(x, y, z)-(2, 3, -\sqrt{3})] = 0$$

o bien

$$x+2y+2\sqrt{3}z-2 = 0.$$

Sea $\mathscr{S}$ una superficie dada en la forma:

$$\mathscr{S} = \{(x, y, z) \mid z = f(x, y), (x, y) \in \mathscr{E}\}$$

donde f es diferenciable sobre el conjunto abierto $\mathscr{E}$ de R^2. Haciendo $F(x, y, z) = f(x, y) - z$, podemos escribir $\mathscr{S}$ en la forma:

$$\mathscr{S} = \{(x, y, z) \mid F(x, y, z) = 0\}$$

donde F es diferenciable sobre el conjunto abierto $\{(x, y, z) \mid (x, y) \in \mathscr{E}, z \in R\}$. Así pues, si $\nabla F(\mathbf{x}_0) \neq \mathbf{0}$, una ecuación del plano tangente a $\mathscr{S}$ en $\mathbf{x}_0 = (x_0, y_0, z_0)$ es

$$(\mathbf{x} - \mathbf{x}_0) \cdot \nabla F(\mathbf{x}_0) = 0.$$

Escribiendo esto en términos de f, obtenemos

12.6 $$(x - x_0) D_1 f(x_0, y_0) + (y - y_0) D_2 f(x_0, y_0) - (z - z_0) = 0$$

como una ecuación del plano tangente a $\mathscr{S}$ en (x_0, y_0, z_0). La ecuación 12.6 puede también escribirse en la forma

$$(x - x_0)\frac{\partial z}{\partial x} + (y - y_0)\frac{\partial z}{\partial y} - (z - z_0) = 0$$

donde las derivadas parciales están evaluadas en el punto (x_0, y_0). Como $\nabla F(x_0, y_0, z_0) = (D_1 f(x_0, y_0), D_2 f(x_0, y_0), -1)$, el gradiente de F no puede ser cero y, por tanto, existe un plano tangente en todo punto de $\mathscr{S}$.

12.7 Ejemplo. Determínese el plano tangente en el punto $(1, 2, 6)$ a la superficie $\mathscr{S}$ de ecuación $z = 2x^2 + y^2$.

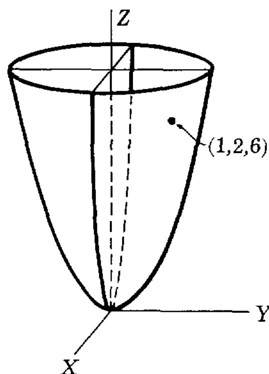


FIGURA 18

Solución. La superficie $\mathscr{S}$ se llama paraboloide elíptico (figura 18). Haciendo $f(x, y) = 2x^2 + y^2$, tenemos

$$D_1 f(x, y) = 4x, \qquad D_2 f(x, y) = 2y$$

y

$$D_1 f(1, 2) = 4, \qquad D_2 f(1, 2) = 4.$$

Por tanto, una ecuación del plano tangente es

$$4(x-1)+4(y-2)-(z-6) = 0$$

o bien

$$4x+4y-z-6 = 0.$$

De la ecuación 12.6 puede deducirse una interpretación geométrica de la diferencial. Sean $(x_0, y_0) = \mathbf{x}$ y $(x, y) = \mathbf{x}$. Entonces $z_0 = f(\mathbf{x}_0)$ y la ecuación 12.6 puede escribirse en la forma

$$\begin{aligned} z &= f(\mathbf{x}_0)+\nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) \\ &= f(\mathbf{x}_0)+df(\mathbf{x}_0;\,\mathbf{x}-\mathbf{x}_0). \end{aligned}$$

Así pues, cuando nos aproximamos al incremento $\Delta f(\mathbf{x}_0;\,\mathbf{x}-\mathbf{x}_0)$ por la diferencial $df(\mathbf{x}_0;\,\mathbf{x}-\mathbf{x}_0)$ para un $|\mathbf{x}-\mathbf{x}_0|$ pequeño, nos estamos aproximando a la superficie por su plano tangente en la vecindad de $\mathbf{x}_0$ (figura 19).

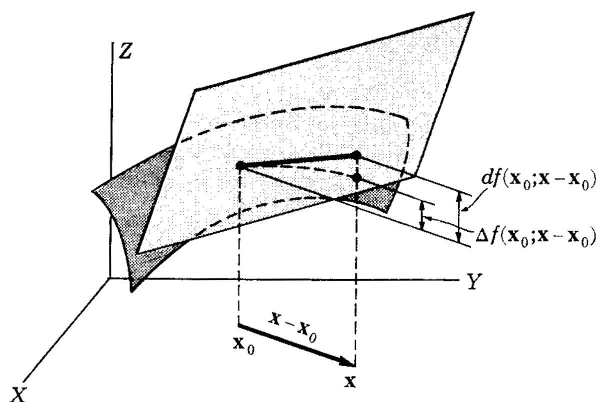


FIGURA 19

Problemas

1. Dibújese la superficie dada y encuéntrese el plano tangente a la superficie en elípunto dado:

a) esfera : $x^2+y^2+z^2 = 16$, $(2, 3, \sqrt{3})$

b) hiperboloide de dos hojas : $\dfrac{x^2}{4} + y^2-z^2 = -1$, $(-2, 3, \sqrt{11})$

c) cilindro elíptico recto : $4x^2+9y^2 = 36$, $\left(\dfrac{3}{4}, \dfrac{\sqrt{15}}{2}, \sqrt{3}\right)$

d) cono circular recto : $x^2+y^2-z^2 = 0$, $(1, 2, \sqrt{5})$.

2. Dibújese el cono elíptico recto de ecuación $4x^2-y^2+z^2 = 0$. ¿Tiene este cono un plano tangente en su vértice $(0, 0, 0)$?

3. Dibújese la superficie dada y encuéntrese el plano tangente a la superficie en el punto dado.

a) plano: $z = x+y$, $(1, 1, 2)$

b) hemisferio: $z = -\sqrt{9-x^2-y^2}$, $(2, 1, -2)$

c) paraboloide hiperbólico: $z = \dfrac{x^2}{4} - \dfrac{y^2}{9}$, $(2, 3, 0)$

d) $z = x^2y^2$, $(1, 2, 4)$.

13. EL TEOREMA DE LA FUNCIÓN IMPLÍCITA

En la sección precedente considerábamos superficie de la forma

13.1 $$\{(x, y, z) \mid z = f(x, y)\}$$

y

13.2 $$\{(x, y, z) \mid F(x, y, z) = 0\}.$$

Mientras que es claro que una superficie que se da en la forma 13.1 puede también representarse en la forma 13.2, la recíproca no es generalmente cierta. Es decir, no toda superficie de la forma 13.2 es la gráfica de una función de R^2 en R. Una superficie no puede ser la gráfica de una función de R^2 en R si tiene más de una intersección con una recta perpendicular al plano XY.

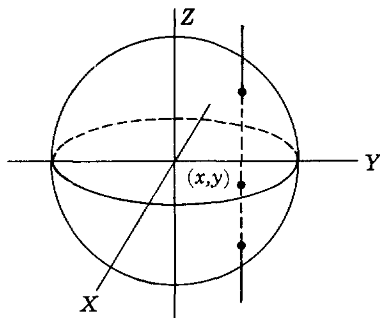


FIGURA 20

Por ejemplo, la esfera

$$\mathscr{S} = \{(x, y, z) \mid x^2+y^2+z^2-4 = 0\}$$

no es la gráfica de una función de R^2 en R, ya que para cualquier punto (x, y)

en el disco $\mathscr{S}((0, 0); 2)$, la recta que pasa por (x, y) perpendicular al plano XY intersecta a la esfera en dos puntos (figura 20).

Si resolvemos la ecuación $x^2+y^2+z^2-4 = 0$ para z, obtenemos $z = \pm\sqrt{4-x^2-y^2}$. Si f_1 y f_2 son las funciones definidas por las reglas

$$f_1(x, y) = \sqrt{4-x^2-y^2}$$
$$f_2(x, y) = -\sqrt{4-x^2-y^2},$$

vemos que la esfera es la unión de las gráficas de estas dos funciones, es decir,

$$\mathscr{S} = \{(x, y, z) \mid z = f_1(x, y)\} \cup \{(x, y, z) \mid z = f_2(x, y)\}.$$

La gráfica de f_1 es el hemisferio superior y la gráfica de f_2 es el hemisferio inferior. En este caso decimos que las funciones f_1 y f_2 están definidas implícitamente por la ecuación $x^2+y^2+z^2-4 = 0$.

En general, decimos que una función f está definida implícitamente por la ecuación

$$F(x, y, z) = 0$$

si, para todo $(x, y) \in \mathscr{D}_f$,

$$F(x, y, f(x, y)) = 0.$$

Podemos escribir esto en términos de conjuntos como sigue: f está definido implícitamente por la ecuación $F(x, y, z) = 0$ si

$$\{(x, y, z) \mid z = f(x, y), (x, y) \in \mathscr{D}_f\} \subset \{(x, y, z) \mid F(x, y, z) = 0\}.$$

Volviendo a una consideración de la esfera

$$\mathscr{S} = \{(x, y, z) \mid x^2+y^2+z^2-4 = 0\},$$

tomemos un punto (x_0, y_0, z_0) sobre el hemisferio superior ($z_0 > 0$). Entonces, para una vecindad suficientemente pequeña $\mathscr{N}$ de (x_0, y_0) en el plano XY, la ecuación de la esfera define implícitamente una función continua única de dominio $\mathscr{N}$ y que contiene (x_0, y_0, z_0) en su gráfica (figura 21). Esta es la función f con dominio $\mathscr{N}$ y regla de correspondencia

$$f(x, y) = \sqrt{4-x^2-y^2}.$$

Si (x_0, y_0, z_0) es un punto del hemisferio inferior ($z_0 < 0$), entonces el mismo resultado se obtiene con f dado por la regla de correspondencia

$$f(x, y) = -\sqrt{4-x^2-y^2}.$$

Sin embargo, si tomamos un punto (x_0, y_0, z_0) sobre la esfera con $z_0 = 0$, digamos el $(0, 2, 0)$, entonces no es posible encontrar una vecindad de $(0, 2)$ en el plano XY para la que haya una función implícita única, continua, que tenga esta vecindad como dominio y que contenga $(0, 2, 0)$ en su

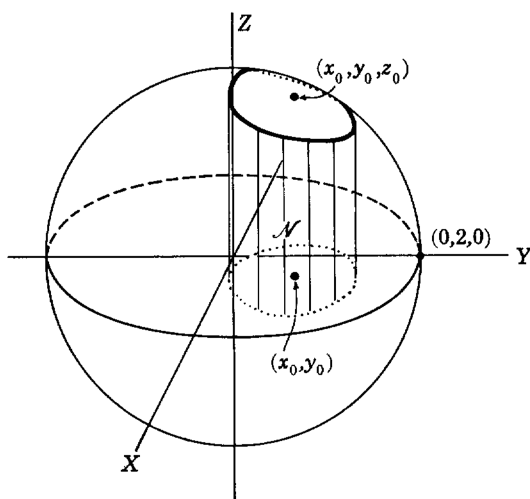


FIGURA 21

gráfica. Los puntos $(x_0, y_0, 0)$ difieren de los otros puntos de la esfera en que en estos puntos el plano tangente a la esfera es paralelo al eje Z.

En el anterior ejemplo fue fácil resolver la ecuación para z en términos de x y y y determinar así explícitamente las funciones definidas implícitamente por la ecuación. Sin embargo, en general esto no es factible. El siguiente teorema afirma la existencia de funciones implícitas bajo ciertas circunstancias y da fórmulas para las derivadas parciales de estas funciones.

13.3 Teorema. (Teorema de la función implícita.) *Sea F una función de* R^3 *en* R *que pertenece a la clase C^1 en un conjunto abierto $\mathscr{E}$. Si $F(\mathbf{x}_0) = 0$ y $D_3 F(\mathbf{x}_0) \neq 0$, donde $\mathbf{x}_0 = (x_0, y_0, z_0) \in \mathscr{E}$, entonces existe una vecindad $\mathcal{N}$ de (x_0, y_0), una vecindad $\langle z_0 - c, z_0 + c\rangle$ de z_0, y una función única $f \in C^1$ sobre $\mathcal{N}$ tal que*

$$f(x_0, y_0) = z_0$$

y, para todo $(x, y) \in \mathcal{N}$,

$$f(x, y) \in \langle z_0 - c, z_0 + c\rangle$$

y

$$F(x, y, f(x, y)) = 0.$$

Además, para $(x, y) \in \mathcal{N}$,

$$D_1 f(x, y) = -\frac{D_1 F(x, y, f(x, y))}{D_3 F(x, y, f(x, y))}$$

y

$$D_2 f(x, y) = -\frac{D_2 F(x, y, f(x, y))}{D_3 F(x, y, f(x, y))}.$$

PRUEBA. Supongamos $D_3 F(\mathbf{x}_0) > 0$. (La prueba es análoga si $D_3 F(\mathbf{x}_0) < 0$.) Como $D_3 F$ es continua en $\mathscr{E}$, $D_3 F(\mathbf{x}) > 0$ para todo $\mathbf{x}$ en cierta vecindad $\mathscr{S}(\mathbf{x}_0; \delta)$ de $\mathbf{x}_0$. Tomemos $0 < c < \delta$. Para x y y fijos, $D_3 F(x, y, z) > 0$ implica que $F(x, y, z)$ aumenta cuando z aumenta. Por tanto, como $F(\mathbf{x}_0) = 0$, tenemos $F(x_0, y_0, z_0 - c) < 0$ y $F(x_0, y_0, z_0 + c) > 0$. La continuidad de F en $\mathscr{E}$ implica la existencia de un número $r \leqslant \sqrt{\delta^2 - c^2}$ tal que, para todo $(x, y) \in \mathscr{S}((x_0, y_0); r)$,

$$F(x, y, z_0 - c) < 0 \quad \text{y} \quad F(x, y, z_0 + c) > 0 \text{ (figura 22).}$$

Sea $\mathscr{N} = \mathscr{S}((x_0, y_0); r)$. Tomemos $(x, y) \in \mathscr{N}$ y sea $g(z) = F(x, y, z)$. Entonces g es continua y creciente en $[z_0 - c, z_0 + c]$. Además, $g(z_0 - c) < 0$ y $g(z_0 + c) > 0$. Luego hay un y sólo un punto $z \in \langle z_0 - c, z_0 + c \rangle$ tal

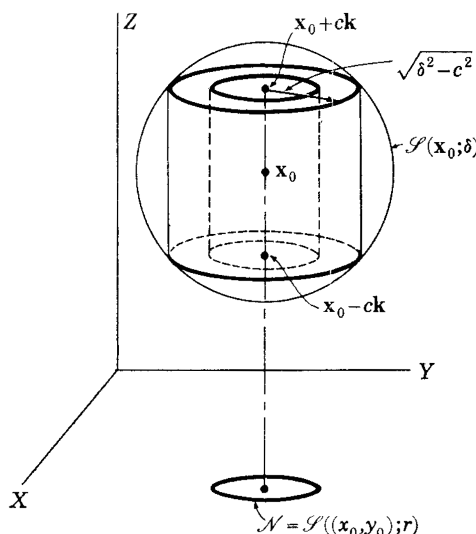


FIGURA 22

que $g(z) = 0$. Así pues, para cada punto $(x, y) \in \mathscr{N}$ existe un z y sólo uno en la vecindad $\langle z_0 - c, z_0 + c \rangle$, tal que $F(x, y, z) = 0$. Si por definición hacemos $f(x, y)$ igual a este z, entonces la función f está definida sobre $\mathscr{N}$ y $F(x, y, f(x, y)) = 0$ para todo $(x, y) \in \mathscr{N}$. Además, $f(x_0, y_0) = z_0$.

Obsérvese que para cualquier $(x, y) \in \mathscr{N}$, $|f(x, y) - f(x_0, y_0)| < c$.

Probamos a continuación que f es continua en $\mathscr{N}$. Tomemos (x, y) en $\mathscr{N}$, sea $z = f(x, y)$, y tomemos ε tal que $0 < \varepsilon < \text{mín}\, \{z_0 + c - z, z - z_0 + c\}$.

Consideremos la región cilíndrica que se extiende de $z-\varepsilon$ a $z+\varepsilon$ y que tiene $\mathcal{N}$ como sección, es decir, como proyección sobre el plano XY. De acuerdo con el argumento usado para demostrar la existencia y unicidad de f podemos comprobar la existencia de una vecindad $\mathcal{N}(x, y)$ de (x, y) tal que para todo punto $(x', y') \in \mathcal{N}(x, y)$

$$|f(x', y') - f(x, y)| < \varepsilon.$$

Por tanto, f es continua en $\mathcal{N}$.

Probaremos ahora que $f \in C^1$ en $\mathcal{N}$. Tomemos $(x, y) \in \mathcal{N}$ y (h, k) tales que $(x+h, y+k) \in \mathcal{N}$. Sea $l(h, k) = f(x+h, y+k) - f(x, y)$. Usando entonces el teorema del valor medio (teorema 6.10, pág. 194), tenemos

$$\begin{aligned} 0 &= F(x+h, y+k, f(x+h, y+k)) - F(x, y, f(x, y)) \\ &= F(x+h, y+k, f(x, y) + l(h, k)) - F(x, y, f(x, y)) \\ &= (h, k, l(h, k)) \cdot \mathbf{D}F(x+\theta h, y+\theta k, f(x, y) + \theta l(h, k)) \text{ donde } \theta \in \langle 0, 1 \rangle. \end{aligned}$$

Si hacemos $k = 0$, obtenemos

$$hD_1 F(x+\theta h, y, f(x, y) + \theta l(h, 0)) \\ + l(h, 0) D_3 F(x+\theta h, y, f(x, y) + \theta l(h, 0)) = 0.$$

Luego, para $h \neq 0$,

$$\frac{l(h, 0)}{h} = \frac{f(x+h, y) - f(x, y)}{h} = -\frac{D_1 F(x+\theta h, y, f(x, y) + \theta l(h, 0))}{D_3 F(x+\theta h, y, f(x, y) + \theta l(h, 0))}.$$

Como f es continua en (x, y) y $D_1 F$ y $D_2 F$ son continuas en $(x, y, f(x, y))$, el límite del segundo miembro cuando h tiende a 0 existe y, por tanto,

$$D_1 f(x, y) = -\frac{D_1 F(x, y, f(x, y))}{D_3 F(x, y, f(x, y))}.$$

De un modo análogo obtenemos

$$D_2 f(x, y) = -\frac{D_2 F(x, y, f(x, y))}{D_3 F(x, y, f(x, y))}.$$

Lo que completa la prueba del teorema.

El teorema 13.3 tiene la siguiente interpretación geométrica: supongamos que se nos da una superficie en la forma $\{(x, y, z) \mid F(x, y, z) = 0\}$ donde $F \in C^1$ en un cierto conjunto abierto y supongamos que $\mathbf{x}_0$ es un punto sobre la superficie en el que el plano tangente no es paralelo al eje $Z (D_3 F(\mathbf{x}_0) \neq 0)$. Entonces, en una vecindad de $\mathbf{x}_0$, la superficie tiene una representación única en la forma $\{(x, y, z) \mid z = f(x, y)\}$.

Según teorema 13.3 puede enunciarse también, usando otra terminología, como sigue:

Supongamos que F pertenece a la clase C^1 en un conjunto abierto $\mathscr{E}$ de R^3 y sea $w = F(x, y, z)$. Si $w = 0$ y $\dfrac{\partial w}{\partial z} \neq 0$ en (x_0, y_0, z_0), la ecuación $F(x, y, z) = 0$ tiene una solución continua única para z en términos de x y y en cierta vecindad de (x_0, y_0) y

$$\frac{\partial z}{\partial x} = -\frac{\dfrac{\partial w}{\partial x}}{\dfrac{\partial w}{\partial z}} \qquad \frac{\partial z}{\partial y} = -\frac{\dfrac{\partial w}{\partial y}}{\dfrac{\partial w}{\partial z}}.$$

Nota. El teorema 13.3 fue enunciado en términos de resolución respecto a la tercera variable, pero es claro que un enunciado análogo también se verifica para el caso en que resolvamos respecto a cualquiera de las otras variables.

13.4 Ejemplo. Demuéstrese que en una vecindad de $(0, 2, -3)$ la ecuación

$$xz^3 + yz + 6 = 0$$

puede resolverse para z en términos de x y y y encuéntrense $\dfrac{\partial z}{\partial x}$ y $\dfrac{\partial z}{\partial y}$ en este punto.

SOLUCIÓN. Sea $w = F(x, y, z) = xz^3 + yz + 6$. Entonces

$$\frac{\partial w}{\partial x} = D_1 F(x, y, z) = z^3$$

$$\frac{\partial w}{\partial y} = D_2 F(x, y, z) = z$$

$$\frac{\partial w}{\partial z} = D_3 F(x, y, z) = 3xz^2 + y.$$

Tenemos pues que $F \in C^1$ sobre R^3 y $D_3 F(0, 2, -3) = 2 \neq 0$. Por tanto, la ecuación puede resolverse para z en términos de x y y en alguna vecindad de $(0, 2, -3)$ y en esta vecindad

$$\frac{\partial z}{\partial x} = -\frac{z^3}{3xz^2 + y} \quad \text{y} \quad \frac{\partial z}{\partial y} = -\frac{z}{3xz^2 + y}.$$

Así pues, en $(0, 2, -3)$, $\dfrac{\partial z}{\partial x} = \dfrac{27}{2}$ y $\dfrac{\partial z}{\partial y} = \dfrac{3}{2}$.

El teorema de la función implícita bidimensional es análogo al teorema 13.3

13.5 Teorema. *Sea F una función de R^2 en R que pertenece a la clase C^1 en un conjunto abierto $\mathscr{E}$. Si $F(\mathbf{x}_0) = 0$ y $D_2F(\mathbf{x}_0) \neq 0$, donde $\mathbf{x}_0 = (x_0, y_0) \in \mathscr{E}$, entonces existe una vecindad $\mathscr{N}$ de x_0, una vecindad $\langle y_0 - c, y_0 + c\rangle$ de y_0, y una función única $f \in C^1$ sobre $\mathscr{N}$ tal que*

$$f(x_0) = y_0$$

y, para todo $x \in \mathscr{N}$, $f(x) \in \langle y_0 - c, y_0 + c\rangle$,

$$F(x, f(x)) = 0,$$

y

$$f'(x) = -\frac{D_1F(x, f(x))}{D_2F(x, f(x))}.$$

La prueba de este teorema se sigue del teorema 13.3 si consideramos F como una función definida sobre R^3, es decir, si hacemos $F(x, y) = G(x, y, z)$ y aplicamos el teorema 13.3 a G (resuelto para y).

La fórmula para f' dada en el teorema 13.5 puede también obtenerse usando la regla de la cadena. Si suponemos que f es una función diferenciable definida sobre algún intervalo abierto $\mathscr{J}$ tal que

$$F(x, f(x)) = 0 \quad \text{y} \quad D_2F(x, f(x)) \neq 0 \quad \text{para toda } x \in \mathscr{J}$$

entonces, haciendo $g(x) = F(x, f(x))$ para $x \in \mathscr{J}$, tenemos

$$\begin{aligned} 0 = g'(x) &= \mathbf{D}F(x, f(x)) \cdot (1, f'(x)) \\ &= D_1F(x, f(x)) + f'(x)D_2F(x, f(x)). \end{aligned}$$

De donde

$$f'(x) = -\frac{D_1F(x, f(x))}{D_2F(x, f(x))} \quad \text{para toda } x \in \mathscr{J}.$$

Consideremos, por ejemplo, la ecuación

$$x^2 + y^2 - 4 = 0.$$

Si suponemos que hay una función diferenciable f definida sobre un intervalo abierto $\mathscr{J}$ tal que

$$x^2 + f^2(x) - 4 = 0, \quad x \in \mathscr{J}$$

entonces, tomando la derivada, obtenemos

$$2x + 2f(x)f'(x) = 0$$

$$f'(x) = -\frac{x}{f(x)}, \qquad x \in \mathscr{J}.$$

En este cálculo no se necesita introducir explícitamente la función f. Supongamos que la ecuación $x^2+y^2-4=0$ define y como una función diferenciable de x sobre algún intervalo abierto y derivemos a continuación. El cálculo toma entonces la forma:

$$x^2+y^2-4=0$$

$$2x+2y\frac{dy}{dx}=0$$

$$\frac{dy}{dx}=-\frac{x}{y}.$$

A esto le llamamos diferenciación implícita. Según el teorema de la funcion implícita podemos ver que esta fórmula para la derivada se verifica en cualquier punto (x, y) sobre la circunferencia de ecuación $x^2+y^2=4$, excepto en los puntos $(2, 0)$ y $(-2, 0)$ donde D_2F es 0.

Problemas

1. Dado el elipsoide definido por la ecuación

$$4x^2+3y^2+z^2-12=0.$$

a) Dibújese la gráfica.

b) Determínese el plano tangente al elipsoide en cualquier punto (x_0, y_0, z_0) de la superficie.

c) Resuélvase la ecuación para z en términos de x y y, y encuéntrense

$$\frac{\partial z}{\partial x} \text{ y } \frac{\partial z}{\partial y}.$$

d) ¿En qué puntos de la superficie es le plano tangente paralelo al eje Z?

e) Usando el teorema 13.3, determínense los puntos sobre la superficie en una vecindad de los cuales la ecuación puede resolverse en forma única para z en términos de x y y y encuéntrense $\dfrac{\partial z}{\partial x}$ y $\dfrac{\partial z}{\partial y}$.

f) ¿En qué puntos de la superficie es el plano tangente paralelo al eje X?

g) Determínense los puntos sobre la superficie en una vecindad de los cuales la ecuación puede resolverse en forma única para x en términos de y y z, y encuéntrense $\dfrac{\partial x}{\partial y}$ y $\dfrac{\partial x}{\partial z}$.

2. ¿Puede resolverse la ecuación $x^2y+\operatorname{sen} yz=0$ en forma única para z en términos de x y y en una vecindad de $(4, 0, 3)$? Demuéstrese que

puede resolverse para y en términos de x y z en una vecindad de ese punto y encuéntrense $\frac{\partial y}{\partial x}$ y $\frac{\partial y}{\partial z}$.

3. Demuéstrese que la ecuación $yz^2+z+3xy = 0$, puede resolverse en forma única para z en términos de x y y en una vecindad de $(0, 0, 0)$. Así pues, en una vecindad de $(0, 0, 0)$, la superficie puede representarse por una ecuación de la forma $z = f(x, y)$ donde $f \in C^1$. Determínese la regla de correspondencia para f.

4. Demuéstrese que en una vecindad del punto que se señala las siguientes ecuaciones pueden resolverse en forma única para z en términos de x y y, y encuéntrense $\frac{\partial z}{\partial x}$ y $\frac{\partial z}{\partial y}$.

a) $x \operatorname{sen} yz-4z+6 = 0$; $(0, 5, \frac{3}{2})$
b) $e^{xyz}+3z = 0$; $(4, 0, -\frac{1}{3})$
c) $\sqrt{x^2+y^2+z^2}-2x^2z^5 = 0$; $(2, \sqrt{59}, 1)$
d) $z^y-x^2z+9y = 0$; $(-3, 2, 3)$.

5. Por diferenciación implícita, encuéntrese $\frac{dy}{dx}$ si

a) $xy^3+3xy-1 = 0$
b) $x \operatorname{sen} xy+y = 0$
c) $xe^y+3y = 0$
d) $\cos y^2+xy = 0$
e) $\operatorname{senh} y+y^2-x = 0$
f) $\arctan \frac{y}{x} - y^2 = 0$

6. Consideremos la ecuación $F(x, y) = y^3-x = 0$.
a) Dibújese la gráfica de esta curva.
b) Resuélvase la ecuación para y en términos de x.
c) Demuéstrese que $D_2F(0, 0) = 0$.
d) ¿Hay algunas discrepancias entre las partes *b* y *c*?

7. Pruébese el teorema 13.5.

14. MÁXIMOS Y MÍNIMOS

En esta sección consideramos el problema de determinar los valores máximo y mínimo relativos de una función real f definida sobre un conjunto abierto $\mathscr{E}$ de $\mathbf{R}^n$. Encontramos primero una condición necesaria: si f tiene un valor máximo o mínimo relativo en $\mathbf{x}_0$ entonces todas las derivadas parciales de orden uno o son nulas o no existen en $\mathbf{x}_0$; podemos también demostrar que en $\mathbf{x}_0$ todas las derivadas direccionales o son iguales a cero o no existen. En cualquier caso la condición no es suficiente para asegurar

la existencia de un valor máximo o mínimo relativo de la función en $\mathbf{x}_0$. Para funciones de R^2 en R obtenemos una condición suficiente para la existencia de un máximo o mínimo relativo que puede extenderse, aunque con dificultad, a funciones definidas sobre espacios de mayor dimensión.

14.1 Definición. *La función f tiene un* ***máximo relativo*** *en el punto $\mathbf{x}_0$ si existe una vecindad $\mathcal{N}$ de $\mathbf{x}_0$ tal que, para todo $\mathbf{x} \in \mathcal{N} \cap \mathcal{D}_f$, $f(\mathbf{x}) \leqslant f(\mathbf{x}_0)$.*

Un *mínimo relativo* de una función se define de modo análogo.

14.2 Definición. *Los* ***valores extremos*** *de una función son los máximos y mínimos relativos de la función.*

Ahora probaremos que los valores extremos de una función definida sobre un conjunto abierto de R^n pueden ocurrir solamente en puntos donde todas las derivadas parciales de primer orden son cero o no existen.

14.3 Teorema. *Si la función f definida sobre un conjunto abierto $\mathscr{E}$ de R^n tiene un valor extremo en $\mathbf{x}_0 \in \mathscr{E}$ y $D_k f(\mathbf{x}_0)$ existe, entonces $D_k f(\mathbf{x}_0) = 0$.*

PRUEBA. Supongamos que f tiene un máximo relativo en $\mathbf{x}_0$. Entonces

$$\lim_{h \to 0^+} \frac{f(\mathbf{x}_0 + h\mathbf{u}_k) - f(\mathbf{x}_0)}{h} \leqslant 0$$

y

$$\lim_{h \to 0^-} \frac{f(\mathbf{x}_0 + h\mathbf{u}_k) - f(\mathbf{x}_0)}{h} \geqslant 0.$$

Como $D_k f(\mathbf{x}_0)$ existe, los dos anteriores límites deben ser iguales a $D_k f(\mathbf{x}_0)$. Luego $D_k f(\mathbf{x}_0) = 0$.

Si f tiene un mínimo relativo en $\mathbf{x}_0$, entonces $-f$ tiene un máximo relativo en $\mathbf{x}_0$. Aplicando la parte del teorema ya probada a $-f$, obtenemos $D_k f(\mathbf{x}_0) = 0$.

Los puntos donde todas las derivadas parciales de f son cero o no existen se llaman *puntos críticos* de f. Así pues, el teorema 14.3 nos dice que los valores extremos de una función definida sobre un conjunto abierto pueden ocurrir solamente en los puntos críticos de la función. Sin embargo, la función no necesariamente tiene un valor extremo en cada uno de sus puntos críticos; demostraremos esto en el ejemplo 14.5.

14.4 Ejemplo. Determínense los valores extremos de la función f definida por $f(x, y) = 2x^2 + 4xy + 5y^2 + 2x - y$.

Solución. Como

$$D_1 f(x, y) = 4x+4y+2$$
$$D_2 f(x, y) = 4x+10y-1,$$

$D_1 f$ y $D_2 f$ existen en todos los puntos de R^2 y, por tanto, el único punto crítico es $(-1, \frac{1}{2})$ donde $D_1 f$ y $D_2 f$ son cero. Por tanto, si f tiene algún valor extremo, tal valor extremo debe ocurrir en el punto $(-1, \frac{1}{2})$ y ser $f(-1, \frac{1}{2}) = -\frac{5}{4}$. Podemos demostrar que $-\frac{5}{4}$ es el valor mínimo de f. Por una rotación de los ejes, podemos escribir[1]

$$2x^2+4xy+5y^2+2x-y = x'^2+6y'^2-\sqrt{5}x'$$
$$= \left(x' - \frac{\sqrt{5}}{2}\right)^2 + 6y'^2 - \frac{5}{4}.$$

Así pues, $f(x, y) \geqslant -\frac{5}{4}$ para todo $(x, y) \in R^2$ y, por tanto, $-\frac{5}{4}$ es el valor mínimo de f. Las curvas de nivel de f aparecen señaladas en la figura 23.

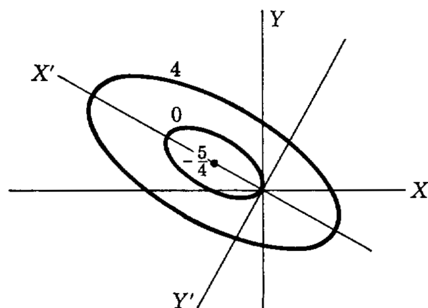


FIGURA 23

14.5 Ejemplo. Determínense cualesquier valores extremos de la función f definida por $f(x, y) = \dfrac{x^2}{3} - \dfrac{3y^2}{16}$.

Solución. Como

$$D_1 f(x, y) = \tfrac{2}{3}x \quad \text{y} \quad D_2 f(x, y) = -\tfrac{3}{8}y,$$

$D_1 f$ y $D_2 f$ existen en todos los puntos de R^2 y son cero solamente en $(0, 0)$. Es pues $(0, 0)$ el único punto crítico de f. Por tanto, si f tiene un valor extremo, tal valor debe ocurrir en el punto $(0, 0)$ y ser $f(0, 0) = 0$. En este

[1] Volumen I, pág. 257.

caso es fácil demostrar que f no tiene un valor extremo en (0, 0). Considerando los valores de f en los puntos sobre el eje X, tenemos

$$f(x, 0) = \tfrac{1}{3}x^2.$$

Así pues, f no puede tener un máximo relativo en (0, 0). Pero considerando los valores de f en los puntos del eje Y, tenemos

$$f(0, y) = -\tfrac{3}{16}y^2.$$

Luego f tampoco puede tener un mínimo relativo en (0, 0). Por tanto, f no tiene un valor extremo en (0, 0). Las curvas de nivel y gráficas de f están dibujadas en las figuras 24a y 24b. El punto (0, 0) se llama *punto de ensilladura* de f. La gráfica de f tiene un plano tangente horizontal en ese punto, pero la función no alcanza en él un valor extremo.

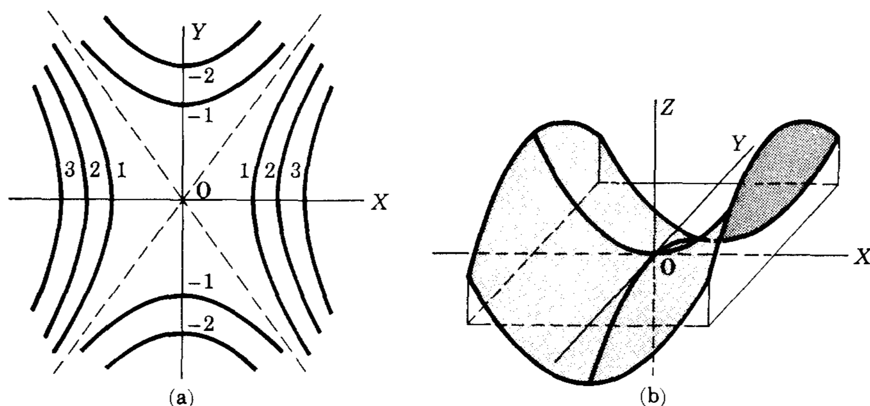


FIGURA 24

El teorema 14.3 nos indica dónde debemos buscar los valores extremos —en los puntos críticos— pero no nos dice cómo saber cuándo nos hemos encontrado con uno. En los ejemplos 14.4 y 14.5 tuvimos que hacer una investigación completa de la función, antes de poder decidir si la función tenía o no un valor extremo en el punto crítico. Sin embargo, hay un teorema, análogo al criterio de la segunda derivada para funciones de R en R, que nos proporciona un método sencillo para conocer cuándo, en un punto crítico, una función de R^2 en R alcanza un valor extremo.

Supongamos que f es una función de R^2 en R que pertenece a la clase C^2 en una vecindad $\mathscr{S}(\mathbf{x}_0; r)$ de $\mathbf{x}_0 = (x_0, y_0)$ y que $\mathbf{D}f(\mathbf{x}_0) = \mathbf{0}$. Entonces, para toda dirección $\mathbf{u}$, $D_\mathbf{u} f(\mathbf{x}_0) = 0$ y, por tanto, la curva formada por la intersección de la superficie $z = f(x, y)$ y el plano vertical que pasa por $(x_0, y_0, 0)$ y es paralelo al vector $(u_1, u_2, 0)$, tiene una tangente horizontal

en el punto $(x_0, y_0, f(x_0, y_0))$ (figura 25). Si f tiene un máximo (mínimo) en $\mathbf{x}_0$, entonces $(x_0, y_0, f(x_0, y_0))$ es un punto máximo (mínimo) de esta curva de intersección. Si $D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0) < 0$ para todo $\mathbf{u}$ donde $D_{\mathbf{u},\mathbf{u}}f = D_{\mathbf{u}}[D_{\mathbf{u}}f]$, entonces, por el criterio de la segunda derivada para máximos y mínimos de funciones de R en R, f tiene un máximo sobre cada una de tales curvas de intersección en $\mathbf{x}_0$. Análogamente, si $D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0) > 0$ para toda $\mathbf{u}$, entonces f tiene un mínimo sobre cada una de tales curvas de intersección, y si $D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0) < 0$ para algún $\mathbf{u}$ y $D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0) > 0$ para algún otro $\mathbf{u}$, f tiene un máximo en $\mathbf{x}_0$ sobre algunas de estas curvas de intersección y un mínimo en $\mathbf{x}_0$ sobre otras. Así pues, es de esperar que si $D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0) < 0$ para toda $\mathbf{u}$, entonces f tenga un máximo en $\mathbf{x}_0$ y si $D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0) > 0$ para todo $\mathbf{u}$, f tenga un mínimo en $\mathbf{x}_0$; mientras que si $D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0)$ es negativa para algún $\mathbf{u}$ y positiva para algún otro, entonces f tenga un punto de ensilladura en $\mathbf{x}_0$.

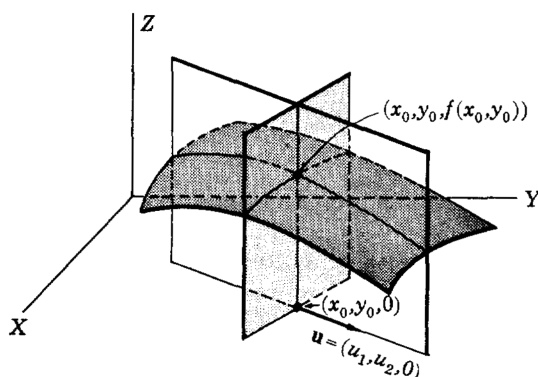


FIGURA 25

Mostraremos que, como esperábamos, es este un criterio para máximos, mínimos y puntos de ensilladura de f, pero, primero, lo transformaremos para ponerlo en términos de derivadas parciales.

$$\begin{aligned} D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0) &= \mathbf{u}\cdot\mathbf{D}[\mathbf{u}\cdot\mathbf{D}f](\mathbf{x}_0) \\ &= u_1^2 D_{1,1}f(\mathbf{x}_0) + 2u_1u_2D_{1,2}f(\mathbf{x}_0) + u_2^2 D_{2,2}f(\mathbf{x}_0) \\ &= au_1^2 + 2bu_1u_2 + cu_2^2 \end{aligned}$$

donde $a = D_{1,1}f(\mathbf{x}_0)$, $b = D_{1,2}f(\mathbf{x}_0)$, y $c = D_{2,2}f(\mathbf{x}_0)$. Si $a \neq 0$, entonces

$$au_1^2 + 2bu_1u_2 + cu_2^2 = a\left[\left(u_1 + \frac{b}{a}u_2\right)^2 + \frac{ac-b^2}{a^2}u_2^2\right].$$

Vemos por esto que si $ac-b^2 > 0$, entonces $D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0)$ es o positivo para todo $\mathbf{u}$ o negativo para todo $\mathbf{u}$, según cuál sea el signo de a. Si $ac-b^2 < 0$, entonces podemos escoger $\mathbf{u}$ de forma que $D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0)$ sea, a voluntad, positivo o negativo. Si $c \neq 0$, los resultados son los mismos. Si $a = c = 0$

y $b \neq 0$, entonces $ac-b^2 < 0$. En este caso $D_{\mathbf{u},\mathbf{u}}f(\mathbf{x}_0) = 2bu_1u_2$, que puede hacerse positivo o negativo con elecciones apropiadas de $\mathbf{u}$.

Ahora enunciaremos y probaremos este criterio.

14.6 Teorema. *Supongamos que f es una función de* R^2 *en* R *que pertenece a la clase C^2 en una vecindad $\mathscr{S}(\mathbf{x}_0; r)$ de $\mathbf{x}_0$, y supongamos que $D_1 f(\mathbf{x}_0) = D_2 f(\mathbf{x}_0) = 0$.*

1. *Si $D_{1,1}f(\mathbf{x}_0)D_{2,2}f(\mathbf{x}_0)-(D_{1,2}f(\mathbf{x}_0))^2 > 0$, entonces f tiene un valor extremo en $\mathbf{x}_0$: un máximo relativo si $D_{1,1}f(\mathbf{x}_0) < 0$ y un mínimo relativo si $D_{1,1}f(\mathbf{x}_0) > 0$.*

2. *Si $D_{1,1}f(\mathbf{x}_0)D_{2,2}f(\mathbf{x}_0)-(D_{1,2}f(\mathbf{x}_0))^2 < 0$, entonces f no tiene un valor extremo en $\mathbf{x}_0$: tiene un punto de ensilladura.*

PRUEBA. Tomemos un punto $\mathbf{x}_0+\mathbf{h}$ en la vecindad reducida $\mathscr{S}'(\mathbf{x}_0; r)$ y sea $\mathbf{h} = (h_1, h_2)$. Usando el teorema de Taylor, pág. 219, obtenemos que para un cierto $\theta \in \langle 0, 1 \rangle$

$$\begin{aligned} f(\mathbf{x}_0+\mathbf{h})-f(\mathbf{x}_0) &= h_1 D_1 f(\mathbf{x}_0)+h_2 D_2 f(\mathbf{x}_0)+\tfrac{1}{2}[h_1{}^2 D_{1,1}f(\mathbf{x}_0+\theta\mathbf{h}) \\ &\qquad +2h_1h_2D_{1,2}f(\mathbf{x}_0+\theta\mathbf{h})+h_2{}^2D_{2,2}f(\mathbf{x}_0+\theta\mathbf{h})] \\ &= \tfrac{1}{2}[h_1{}^2 D_{1,1}f(\mathbf{x}_0+\theta\mathbf{h})+2h_1h_2D_{1,2}f(\mathbf{x}_0+\theta\mathbf{h}) \\ &\qquad +h_2{}^2D_{2,2}f(\mathbf{x}_0+\theta\mathbf{h})] \\ &= \tfrac{1}{2}[Ah_1{}^2+2Bh_1h_2+Ch_2{}^2] \end{aligned}$$

donde $A = D_{1,1}f(\mathbf{x}_0+\theta\mathbf{h})$, $B = D_{1,2}f(\mathbf{x}_0+\theta\mathbf{h})$, $C = D_{2,2}f(\mathbf{x}_0+\theta\mathbf{h})$. Sea $g = D_{1,1}fD_{2,2}f-(D_{1,2}f)^2$; entonces g es continua sobre $\mathscr{S}(\mathbf{x}_0; r)$.

1. Si $g(\mathbf{x}_0) > 0$ y $D_{1,1}f(\mathbf{x}_0) < 0$, entonces existe una vecindad $\mathscr{S}(\mathbf{x}_0;\delta) \subset \mathscr{S}(\mathbf{x}_0; r)$ tal que para todo $\mathbf{x}\in\mathscr{S}(\mathbf{x}_0; \delta)$

$$g(\mathbf{x}) > 0 \quad \text{y} \quad D_{1,1}f(\mathbf{x}) < 0.$$

Si tomamos $\mathbf{x}_0+\mathbf{h}\in\mathscr{S}(\mathbf{x}_0; \delta)$, entonces $\mathbf{x}_0+\theta\mathbf{h}\in\mathscr{S}(\mathbf{x}_0; \delta)$ y

$$g(\mathbf{x}_0+\theta\mathbf{h}) = AC-B^2 > 0 \quad \text{y} \quad D_{1,1}f(\mathbf{x}_0+\theta\mathbf{h}) = A < 0.$$

Por tanto, para todo $\mathbf{x}_0+\mathbf{h}\in\mathscr{S}'(\mathbf{x}_0; \delta)$,

$$\begin{aligned} f(\mathbf{x}_0+\mathbf{h})-f(\mathbf{x}_0) &= \tfrac{1}{2}[Ah_1{}^2+2Bh_1h_2+Ch_2{}^2] \\ &= \frac{1}{2A}[(Ah_1+Bh_2)^2+(AC-B^2)h_2{}^2] \\ &< 0. \end{aligned}$$

Esto prueba que f tiene un máximo relativo en $\mathbf{x}_0$ si $g(\mathbf{x}_0) > 0$ y $D_{1,1}f(\mathbf{x}_0) < 0$. La prueba que f tiene un mínimo relativo en $\mathbf{x}_0$ si $g(\mathbf{x}_0) > 0$ y $D_{1,1}f(\mathbf{x}_0) > 0$ es análoga a la anterior.

2. Si $g(\mathbf{x}_0) < 0$, demostraremos que hay dos rectas $\mathscr{L}_1$ y $\mathscr{L}_2$ que pasan por $\mathbf{x}_0$ tales que $f(\mathbf{x}_0)$ es un mínimo relativo de los valores de la función sobre una de las rectas y es un máximo relativo de los valores de la función sobre la otra recta.

Si $\mathbf{h} = |\mathbf{h}|\,\mathbf{u} = |\mathbf{h}|(u_1, u_2)$, entonces

$$f(\mathbf{x}_0+\mathbf{h})-f(\mathbf{x}_0) = \tfrac{1}{2}|\mathbf{h}|^2\,[Au_1{}^2+2Bu_1u_2+Cu_2{}^2].$$

Sea $a = D_{1,1}f(\mathbf{x}_0)$, $b = D_{1,2}f(\mathbf{x}_0)$, $c = D_{2,2}f(\mathbf{x}_0)$. Como $f \in C^2$ sobre $\mathscr{S}(\mathbf{x}_0; r)$, para toda $|\mathbf{h}|$ suficientemente pequeña $Au_1{}^2+2Bu_1u_2+Cu_2{}^2$, y por tanto $f(\mathbf{x}_0+\mathbf{h})-f(\mathbf{x}_0)$, tiene el mismo signo que $au_1{}^2+2bu_1u_2+cu_2{}^2$, con tal que este último sea realmente distinto de cero. Ahora demostraremos que siempre hay dos elecciones de $\mathbf{u} = (u_1, u_2)$ tales que $au_1{}^2+2bu_1u_2+cu_2{}^2$ tiene signos opuestos para estas dos elecciones. Es decir, tales que $f(\mathbf{x}_0)$ es un mínimo relativo para una de estas dos elecciones y un máximo relativo para la otra.

Consideraremos tres casos.

Caso 1. $a \neq 0$.

Si $(u_1, u_2) = (1, 0)$, entonces $au_1{}^2+2bu_1u_2+cu_2{}^2 = a$.

Si $(u_1, u_2) = \dfrac{1}{\sqrt{a^2+b^2}}(b, -a)$, entonces

$$au_1{}^2+2bu_1u_2+cu_2{}^2 = \frac{a}{a^2+b^2}\,g(\mathbf{x}_0)\,.$$

Así pues $g(\mathbf{x}_0) < 0$, para $|\mathbf{h}|$ suficientemente pequeño, pero no cero, el signo de $f(\mathbf{x}_0+\mathbf{h})-f(\mathbf{x}_0)$ será diferente cuando $\mathbf{x}_0+\mathbf{h} \in \mathscr{L}_1 = \{\mathbf{x}_0+t(1, 0)\}$ que cuando $\mathbf{x}_0+\mathbf{h} \in \mathscr{L}_2 = \{\mathbf{x}_0+t(b, -a)\}$. Así pues $f(\mathbf{x}_0)$ es un valor mínimo relativo sobre una de las rectas y un valor máximo relativo sobre la otra.

Caso 2. $c \neq 0$.

Este caso es análogo al caso 1. Aquí $f(\mathbf{x}_0)$ es un valor mínimo relativo respecto a una de la rectas, $\mathscr{L}_1 = \{\mathbf{x}_0+t(0, 1)\}$ y $\mathscr{L}_2 = \{\mathbf{x}_0+t(c, -b)\}$, y es un máximo relativo sobre la otra recta.

Caso 3. $a = c = 0$.

Como $g(\mathbf{x}_0) = ac-b^2 < 0$, debemos tener $b \neq 0$.

Si $(u_1, u_2) = \dfrac{1}{\sqrt{2}}(1, 1)$, entonces $au_1{}^2+2bu_1u_2+cu_2{}^2 = b$.

Si $(u_1, u_2) = \dfrac{1}{\sqrt{2}}(1, -1)$, entonces $au_1{}^2+2bu_1u_2+cu_2{}^2 = -b$.

Así pues, para $|\mathbf{h}|$ suficientemente pequeño, pero no cero, el signo de $f(\mathbf{x}_0+\mathbf{h})-f(\mathbf{x}_0)$ será diferente cuando $\mathbf{x}_0+\mathbf{h}\in\mathscr{L}_1 = \{\mathbf{x}_0+t(1, 1)\}$ que cuando $\mathbf{x}_0+\mathbf{h}\in\mathscr{L}_2 = \{\mathbf{x}_0+t(1, -1)\}$. Y de nuevo $f(\mathbf{x}_0)$ es un valor mínimo relativo sobre una de las rectas y un valor máximo relativo sobre la otra. Y esto completa la prueba.

Aplicamos ahora el teorema 14.6 a las funciones consideradas en los ejemplos 14.4 y 14.5.

14.7 Ejemplo. ¿Tiene la función f definida por

$$f(x, y) = 2x^2+4xy+5y^2+2x-y$$

un valor extremo en su punto crítico $(-1, \frac{1}{2})$?

SOLUCIÓN. Las derivadas parciales primeras y segundas de f son

$$D_1 f(x, y) = 4x+4y+2 \qquad D_2 f(x, y) = 4x+10y-1$$
$$D_{1,1} f(x, y) = 4 \qquad D_{1,2} f(x, y) = 4 \qquad D_{2,2} f(x, y) = 10.$$

Como $D_1 f(-1, \frac{1}{2}) = 0$, $D_2 f(-1, \frac{1}{2}) = 0$, y

$$D_{1,1} f(-1, \tfrac{1}{2}) D_{2,2} f(-1, \tfrac{1}{2}) - (D_{1,2} f(-1, \tfrac{1}{2}))^2 = 24 > 0,$$

f tiene un valor extremo en $(-1, \frac{1}{2})$. Para ser precisos, f tiene un mínimo relativo en $(-1, \frac{1}{2})$, ya que $D_{1,1} f(-1, \frac{1}{2}) > 0$.

14.8 Ejemplo. ¿Tiene la función f definida por

$$f(x, y) = \frac{x^2}{3} - \frac{3y^2}{16}$$

un valor extremo en su punto crítico $(0, 0)$?

SOLUCIÓN. Las derivadas parciales primeras y segundas de f son:

$$D_1 f(x, y) = \tfrac{2}{3}x \qquad D_2 f(x, y) = -\tfrac{3}{8}y$$
$$D_{1,1} f(x, y) = \tfrac{2}{3} \qquad D_{1,2} f(x, y) = 0 \qquad D_{2,2} f(x, y) = -\tfrac{3}{8}.$$

Como $D_1 f(0, 0) = 0$, $D_2 f(0, 0) = 0$, y

$$D_{1,1} f(0, 0) D_{2,2} f(0, 0) - (D_{1,2} f(0, 0))^2 = -\tfrac{1}{4} < 0,$$

f no tiene un valor extremo en $(0, 0)$; f tiene en $(0, 0)$ un punto de ensilladura.

En muchos problemas buscamos los valores extremos de una función sujeta a ciertas restricciones. Por ejemplo, podemos necesitar encontrar los valores extremos de una función F de R^3 en R para puntos sobre una superficie $G(x, y, z) = 0$. La ecuación $G(x, y, z) = 0$ se llama *restricción* o *ecuación de enlace*. Un posible método para manejar tales problemas es el

de resolver la ecuación de enlace para una de las variables en términos de las otras, por ejemplo $z = g(x, y)$, y luego hacer $f(x, y) = F(x, y, g(x, y))$, para encontrar finalmente los valores extremos de f.

14.9 Ejemplo. Una caja rectangular sin tapa ha de tener una superficie de área S. Encuéntrense las dimensiones de la caja de máximo volumen.

SOLUCIÓN. Supongamos que la caja tiene las dimensiones x, y y z, donde z es la altura. Entonces el volumen es xyz y el área de la superficie es $2xz+2yz+xy = S$. Deseamos, pues, encontrar el valor máximo de la función $F(x, y, z) = xyz$ sujeta a la restricción $2xz+2yz+xy-S = 0$. Resolviendo la ecuación de enlace para z, obtenemos

$$z = \frac{S-xy}{2x+2y}.$$

Sea $f(x, y) = xy\dfrac{S-xy}{2x+2y}$. Entonces

$$\begin{aligned} D_1 f(x, y) &= xy\frac{-2xy-2y^2-2S+2xy}{(2x+2y)^2} + y\frac{S-xy}{2x+2y} \\ &= \frac{2y^2}{(2x+2y)^2}(-2xy+S-x^2) \end{aligned}$$

y

$$\begin{aligned} D_2 f(x, y) &= xy\frac{-2x \;\; -2xy-2S+2xy}{(2x+2y)^2} + x\frac{S-xy}{2x+2y} \\ &= \frac{2x^2}{(2x+2y)^2}(-2xy+S-y^2). \end{aligned}$$

Como en este problema x y y deben ser positivas, $D_1 f$ y $D_2 f$ existen en todos los puntos y son cero si y sólo si

$$\begin{aligned} x^2+2xy-S &= 0 \\ y^2+2xy-S &= 0. \end{aligned}$$

Restando la segunda ecuación de la primera, obtenemos $x^2-y^2 = 0$ y, por tanto, $x = y$. Sustituyendo entonces en la primera ecuación, tenemos $3x^2-S = 0$ y, por tanto,

$$x = y = \sqrt{\frac{S}{3}}.$$

Así pues, para x y y positivos, la función f tiene un valor extremo solamente en $\left(\sqrt{\frac{S}{3}}, \sqrt{\frac{S}{3}}\right)$. Debido a la naturaleza del problema este debe ser le valor máximo deseado. Por tanto, la caja tiene volumen máximo cuando

$$x = \sqrt{\frac{S}{3}}, \quad y = \sqrt{\frac{S}{3}}, \quad \text{y} \quad z = \frac{1}{2}\sqrt{\frac{S}{3}}.$$

En el próximo capítulo discutiremos otro método para el manejo de problemas extremos con restricciones en el que no resolvemos explícitamente la ecuación de enlace para una variable en términos de las otras. Este nuevo método se llama el de los multiplicadores de Lagrange.

Problemas

1. Encuéntrense los valores extremos de las siguientes funciones, primero sin usar el cálculo y después usándolo. Dibújese en cada caso un diagrama de curvas de nivel y la gráfica.

a) $f(x, y) = x^2+y^2$ *b*) $f(x, y) = \sqrt{x^2+y^2}$
c) $f(x, y) = x^2+y^2+2x-3y+4$ *d*) $f(x, y) = \text{sen } xy.$

2. Encuéntrense los puntos críticos de las siguientes funciones e identifíqueseles como máximos relativos, mínimos relativos o puntos de ensilladura.

a) $f(x, y) = x^3+3x^2-y^2+4$
b) $f(x, y) = x^2-y^2+2x-3y+4$
c) $f(x, y) = e^{xy}$
d) $f(x, y) = x^4-3x^2-y^2+12$
e) $f(x, y) = \text{sen } x \text{ sen } y$
f) $f(x, y) = x^2+3xy+2y^2-x+3$
g) $f(x, y) = x^2y^2+x^2-4y^2+5x-3$
h) $f(x, y) = x^4-y^4-2x^2+y^2+5.$

3. Encuéntrese la distancia más corta entre las rectas definidas por

$$\mathbf{P} = (1, 3, 5)+s(2, 0, -1) \quad \text{y} \quad \mathbf{P} = (2, 8, 11)+t(-4, 3, 1);$$

Encuéntrense los puntos sobre las rectas que están entre sí a tal distancia y demuéstrese que la recta que pasa por estos puntos es ortogonal a las dos rectas dadas.

4. Encuéntrese la distancia más corta del punto $(2, -3, 1)$ al plano de ecuación $z = 2x+5y-3$.

5. Si $f(x, y) = (y-x^2)(y-2x^2)$ demuéstrese que sobre cada recta que pasa por el origen f tiene un mínimo relativo en el origen. Dibújense

las parábolas $y = x^2$ y $y = 2x^2$ e indíquese en qué regiones del plano f tiene valores positivos y en cuáles negativos. Dedúzcase de esto que f no tiene un valor extremo en (0, 0).

6. Encuéntrense los valores extremos de las siguientes funciones sujetas a las restricciones que se señalan.

a) $f(x, y, z) = x^2+y^2-z$, $2x-3y+z-6 = 0$
b) $f(x, y, z) = ze^{-xy}$, $x^2+y^2-z = 0$
c) $f(x, y, z) = x^2+y^2+z^2-2x+1$, $\sqrt{x^2+y^2}-z = 0$.

7. Determínese el paralelepípedo rectangular de máximo volumen y área total igual a 48.

8. Determínese el paralelepípedo rectangular de máximo volumen y lados paralelos a los planos coordenados que puede inscribirse en el elipsoide

$$\frac{x^2}{2} + \frac{y^2}{5} + \frac{z^2}{4} = 1.$$

15. RESUMEN

En este capítulo hemos presentado el cálculo diferencial de funciones reales de un vector (también llamadas funciones reales de más de una variable). La teoría para funciones de una variable vectorial, aunque más complicada en detalle, es una extensión natural de la teoría de funciones de una variable real. La definición de límite de una función de un vector es formalmente la misma que la de límite para una función de variable real. La extensión de la derivada se efectúa definiendo la noción de diferenciabilidad para definir la derivada en términos de esta noción.

Las derivadas direccionales y las derivadas parciales también fueron otros conceptos definidos y se estudiaron las relaciones entre estos distintos tipos de derivadas. Las derivadas parciales son derivadas direccionales en direcciones particulares. Si f es diferenciable en un punto $\mathbf{x}$ entonces todas las derivadas direccionales de f existen en ese punto y el valor de la derivada es un vector cuya dirección es aquella en que la derivada direccional tiene su valor máximo, y cuya longitud es el valor de esta derivada direccional máxima. Además, la derivada de f en $\mathbf{x}$ es el gradiente de f en $\mathbf{x}$: un vector cuyos componentes son las derivadas parciales de f en $\mathbf{x}$.

El cálculo de las derivadas parciales de una función puede hacerse calculando la derivada de una función real de una variable real. Por tanto, como los valores de la derivada y de las derivadas direccionales de una función diferenciable pueden obtenerse partiendo de los valores de las derivadas parciales, no son necesarias, realmente, ningunas nuevas técnicas

de cálculo. Además, la consideración de las derivadas parciales nos proporciona una condición suficientemente sencilla de diferenciabilidad: si una función tiene derivadas parciales continuas, es diferenciable.

Problemas de repaso

1. Verifíquese, usando la definición de límite, que

$$\lim_{(x,y)\to(1,2)} (x^2+y) = 3$$

2. Sea $f(x, y) = \dfrac{x^3}{xy+y}$

a) Demuéstrese que $\lim\limits_{(x,y)\to(0,0)} f(x, y) = 0$ si $(x, y)\in\{(x, y) \mid y = mx\}$, donde $m \neq 0$.

b) Demuéstrese que $\lim\limits_{(x,y)\to(0,0)} f(x, y) = 1$ si $(x, y)\in\{(x, y) \mid y = x^3\}$.

c) ¿Existe $\lim\limits_{(0,0)} f$?

3. Existe $\lim\limits_{(x,y)\to(1,2)} \dfrac{xy-2x-y+2}{x^2+y^2-2x-4y+5}$?

4. Sea $f(x, y) = \dfrac{y}{y+x^2}$. Determínese $\lim\limits_{x\to 0}\lim\limits_{y\to 0} f(x, y)$ y $\lim\limits_{y\to 0}\lim\limits_{x\to 0} f(x, y)$

¿Qué puede decirse de $\lim\limits_{(x,y)\to(0,0)} f(x, y)$?

5. Si $f(x, y, z) = x^2 \operatorname{sen}(yz)$, determínense las derivadas parciales de f. Determínese, ademas, $D_{\mathbf{u}}f$ donde $\mathbf{u} = \dfrac{1}{\sqrt{6}}(-1, 2, 1)$.

6. Si $f(x, y) = \dfrac{xy}{x^2+y^2}$ para $(x, y) \neq (0, 0)$ y $f(0, 0) = 0$, determínense $D_1 f(0, 0)$ y $D_2 f(0, 0)$.

7. Encuéntrese la dirección y magnitud de la máxima razón de cambio de la función f definida por $f(x, y, z) = 3x^4y-yz^2$, en el punto $(0, 3, 1)$.

8. Si las derivadas parciales de las funciones $f_i (i = 1, \ldots, m)$ con respecto a la coordenada k-ésima existen sobre un conjunto abierto $\mathscr{E}$, demuéstrese que

a) $\sum D_k \overset{m}{\underset{i=1}{}} f_i = \sum\limits_{i=1}^{m} D_k f_i$

b) $D_k \prod_{i=1}^{m} f_i = \sum_{j=1}^{m} \frac{D_k f_j}{f_j} \prod_{i=1}^{m} f_i.$

9. Si la función f de R^2 en R está definida sobre una vecindad $\mathscr{S}((a, b); r)$ y si $D_2 f(x, y) = 0$ para todo $(x, y) \in \mathscr{S}((a, b); r)$, demuéstrese que podemos considerar f como una función de una sola variable; es decir, demuéstrese que podemos definir una función g de R en R por la regla $g(x) = f(x, y)$ sobre $\mathscr{S}((a, b); r)$.

10. Desarróllese la fórmula de Taylor con (0, 0) como punto inicial para la función f definida por $f(x, y) = e^x \operatorname{sen} y$, incluyendo todos los términos de tercer grado.

11. Determínese la ecuación del plano tangente al cilindro circular recto de ecuación $x^2 + y^2 = 9$ en el punto (0, 3, 2).

12. Determínese la ecuación del plano tangente al cono elíptico recto de ecuación $z = \sqrt{x^2 + 2y^2}$ en el punto (2, 0, 2).

13. Encuéntrense los valores extremos de la función f definida por $f(x, y) = x^2 - 3xy + 3y^2 - 4x + 5$.

14. Un punto $\mathbf{x}_0$ se llama *punto aislado* de $\mathscr{E}$ si $\mathbf{x}_0 \in \mathscr{E}$ y hay una vecindad reducida de $\mathbf{x}_0$ contenida en $\mathscr{C}\mathscr{E}$. Demuéstrese que todo punto aislado de $\mathscr{E}$ es un punto frontera de $\mathscr{E}$.

15. Proporciónense $\mathscr{E}_i$, $\mathscr{E}_b$, $\mathscr{E}_e$, $\bar{\mathscr{E}}$, y el conjunto de puntos aislados de $\mathscr{E}$ cuando

a) $\mathscr{E} = \{(x, y) \mid 0 < x^2 + y^2 < 16\}$
b) $\mathscr{E} = \{(x, y) \mid x^4 - y^4 - 4x^2 - 4y^2 = 0\}$
c) $\mathscr{E} = \{(x, y, z) \mid z > x^2 + y^2\}$
d) $\mathscr{E} = \left\{(x, y) \mid y = \operatorname{sen} \frac{1}{x}\right\}$.





Funciones vectoriales de un vector

1. INTRODUCCIÓN

Una *función vectorial* **f** *de un vector* es una correspondencia desde un conjunto $\mathscr{A}$ de vectores a un conjunto $\mathscr{B}$ de vectores tal que para cada vector $\mathbf{a} \in \mathscr{A}$ hay un y sólo un vector correspondiente $\mathbf{f}(\mathbf{a}) \in \mathscr{B}$; es decir, es una transformación del conjunto $\mathscr{A}$ en el conjunto $\mathscr{B}$. Si $\mathscr{A}$ es un conjunto en R^n y $\mathscr{B}$ es un conjunto en R^m, entonces decimos que **f** es una función de R^n en R^m. Las funciones de R en R^n y de R^n en R que consideramos en los dos capítulos anteriores son casos especiales de este tipo de funciones.

En geometría analítica se consideran transformaciones del plano; estas son funciones vectoriales de un vector con dominio y rango en R^2. Por ejemplo, la función **f** definida por

$$\mathbf{f}(x, y) = (x, y) + (2, 3) = (x+2, y+3)$$

es una traslación del plano; cada punto (x, y) en R^2 se transforma en el punto $(x+2, y+3)$ en R^2.

El gradiente de una función de R^n en R es una función de R^n en R^n. Por ejemplo, si

$$f(x, y, z) = x^2 y + yz$$

entonces

$$\nabla f(x, y, z) = (2xy, x^2+z, y);$$

es decir, si

$$f = I_1{}^2 I_2 + I_2 I_3$$

entonces

$$\nabla f = (2I_1 I_2, I_1{}^2 + I_3, I_2).$$

Las funciones $2I_1 I_2$, $I_1{}^2 + I_3$, e I_2 son funciones de R^3 en R y se llaman funciones componentes de ∇f.

En general, si $\mathbf{f}$ es una función de R^n en R^m, entonces escribimos $\mathbf{f} = (f_1, \ldots, f_m)$ donde la función componente $f_k (k = 1, \ldots, m)$ es la función de R^n en R con dominio $\mathscr{D}_{\mathbf{f}}$ y regla de correspondencia: $f_k(\mathbf{x})$ es el k-ésimo componente del vector $\mathbf{f}(\mathbf{x})$. Veremos que, como en el caso de funciones de R en R^n (capítulo 3), el cálculo de funciones de R^n en R^m puede expresarse en términos de las funciones componentes que en este caso son funciones de R^n en R.

En los problemas físicos, donde usualmente n es 2, 3 o 4 y m es 2 o 3, las funciones de R^n en R^m se llaman a menudo campos vectoriales. Un ejemplo de una función de R^3 en R^3 es el campo de velocidades de una corriente estacionaria (es decir, con velocidad independiente del tiempo) de un fluido. A cada punto $\mathbf{x}$ del fluido corresponde un vector $\mathbf{v}(\mathbf{x})$: la velocidad de una partícula en el punto $\mathbf{x}$. Una representación geométrica de la función $\mathbf{v}$ puede obtenerse dibujando en el punto $\mathbf{x}$ una flecha que represente el vector $\mathbf{v}(\mathbf{x})$. Si la corriente del fluido no es fija, es decir, si depende del tiempo, entonces la velocidad es una función de R^4 en R^3: $\mathbf{v}(x, y, z, t)$ es la velocidad de una partícula en el punto (x, y, z) en el instante t.

Problema

Proporciónese una imagen geométrica de la función $\mathbf{f}$ dibujando una flecha que represente $\mathbf{f}(\mathbf{x})$ en el punto $\mathbf{x}$ cuando

a) $\mathbf{f}(x, y) = -\dfrac{1}{\sqrt{x^2+y^2}}(x, y)$ *b*) $\mathbf{f}(x, y) = (-y, x)$.

2. LÍMITE Y CONTINUIDAD

Como es de esperar, dadas nuestras experiencias previas con límites, $\lim_{a} \mathbf{f} = \mathbf{b}$ quiere decir que $\mathbf{f}(\mathbf{x})$ está próximo a $\mathbf{b}$ cuando $\mathbf{x}$ está próximo a, pero es distinto de, $\mathbf{a}$. También como de costumbre, suponemos que $\mathbf{a}$ es un punto de acumulación del dominio de $\mathbf{f}$. Definimos entonces el límite como sigue.

2.1 Definición. *Se dice que el vector* $\mathbf{b}$ *es el* ***límite de la función*** $\mathbf{f}$ ***en*** $\mathbf{a}$, *y escribimos* $\lim_{\mathbf{a}} \mathbf{f} = \mathbf{b}$ o $\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{b}$, *si para cada número* $\varepsilon > 0$, *hay un número* $\delta > 0$, *tal que siempre que* $\mathbf{x}$ *esté en el dominio de* $\mathbf{f}$ *y* $0 < |\mathbf{x}-\mathbf{a}| < \delta$ *entonces*

$$|\mathbf{f}(\mathbf{x})-\mathbf{b}| < \varepsilon.$$

Así pues, decimos que $\lim_{\mathbf{a}} \mathbf{f} = \mathbf{b}$ si para cada vecindad $\mathscr{S}(\mathbf{b}; \varepsilon)$ de $\mathbf{b}$ hay una vecindad reducida $\mathscr{S}'(\mathbf{a}; \delta)$ de $\mathbf{a}$ tal que $\mathbf{f}(\mathbf{x}) \in \mathscr{S}(\mathbf{b}; \varepsilon)$ siempre que $\mathbf{x} \in \mathscr{D}_{\mathbf{f}} \cap \mathscr{S}'(\mathbf{a}; \delta)$ (figura 1).

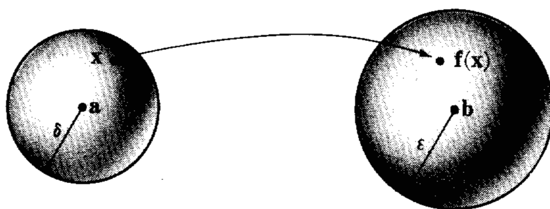


FIGURA 1

La relación entre el límite de una función vectorial de un vector y los límites de sus funciones componentes está dada en el siguiente teorema.

2.2 Teorema. *Sea* $\mathbf{b} = (b_1, \ldots, b_m) \in \mathrm{R}^m$, $\mathbf{f} = (f_1, \ldots, f_m)$ *una función de* R^n *en* R^m, *y* $\mathbf{a}$ *un punto de acumulación de* $\mathscr{D}_{\mathbf{f}}$. *Entonces* $\lim_{\mathbf{a}} \mathbf{f} = \mathbf{b}$ *si y sólo si* $\lim_{\mathbf{a}} f_k = b_k$ *para cada* $k = 1, \ldots, m$.

Omitimos la prueba de este teorema ya que es la misma que la prueba del teorema correspondiente para funciones de R en R^n (pág. 00).

Como un resultado del teorema 2.2, los problemas respecto al límite de una función de R^n en R^m pueden reducirse a problemas sobre los límites de funciones de R^n en R.

Para funciones de R^n en R^m definimos en la forma usual las operaciones de adición, sustracción, multiplicación por un escalar, producto escalar y producto vectorial (solamente si $m = 3$) (véase pág. 104); por ejemplo, $\mathbf{f}+\mathbf{g}$ es la función con dominio $\mathscr{D}_{\mathbf{f}} \cap \mathscr{D}_{\mathbf{g}}$ y regla de correspondencia

$[\mathbf{f}+\mathbf{g}](\mathbf{x}) = \mathbf{f}(\mathbf{x})+\mathbf{g}(\mathbf{x})$. Partiendo de la definición de estas operaciones es fácil demostrar que si $\mathbf{f} = (f_1, \ldots, f_m)$ y $\mathbf{g} = (g_1, \ldots, g_m)$ entonces

$$\mathbf{f}+\mathbf{g} = (f_1+g_1, \ldots, f_m+g_m)$$

$$\mathbf{f}-\mathbf{g} = (f_1-g_1, \ldots, f_m-g_m)$$

$$\varphi \mathbf{f} = (\varphi f_1, \ldots, \varphi f_m)$$

$$\mathbf{f}\cdot\mathbf{g} = \sum_{k=1}^{m} f_k g_k$$

$$\mathbf{f}\times\mathbf{g} = (f_2 g_3 - f_3 g_2, f_3 g_1 - f_1 g_3, f_1 g_2 - f_2 g_1).$$

Usando el teorema 2.2 y el teorema del límite para operaciones sobre funciones de R^n en R (teorema 4.5, pág. 176), obtenemos el siguiente teorema.

2.3 Teorema. *Si* $\mathbf{f}$ *y* $\mathbf{g}$ *son funciones de* R^n *en* R^m *tales que* $\lim_{\mathbf{a}} \mathbf{f}$ *y* $\lim_{\mathbf{a}} \mathbf{g}$ *existen y si* $\mathbf{a}$ *es un punto de acumulación de* $\mathscr{D}_\mathbf{f} \cap \mathscr{D}_\mathbf{g}$, *entonces*

$$\lim_{\mathbf{a}} (\mathbf{f}+\mathbf{g}) = \lim_{\mathbf{a}} \mathbf{f} + \lim_{\mathbf{a}} \mathbf{g}$$

$$\lim_{\mathbf{a}} (\mathbf{f}-\mathbf{g}) = \lim_{\mathbf{a}} \mathbf{f} - \lim_{\mathbf{a}} \mathbf{g}$$

$$\lim_{\mathbf{a}} (\mathbf{f}\cdot\mathbf{g}) = (\lim_{\mathbf{a}} \mathbf{f})\cdot(\lim_{\mathbf{a}} \mathbf{g})$$

$$\lim_{\mathbf{a}} (\mathbf{f}\times\mathbf{g}) = (\lim_{\mathbf{a}} \mathbf{f})\times(\lim_{\mathbf{a}} \mathbf{g}) \qquad [m = 3].$$

Además, si φ *es una función de* R^n *en* R *y* $\mathbf{a}$ *es un punto de acumulación de* $\mathscr{D}_\mathbf{f} \cap \mathscr{D}_\varphi$, *entonces*

$$\lim_{\mathbf{a}} (\varphi\mathbf{f}) = (\lim_{\mathbf{a}} \varphi)(\lim_{\mathbf{a}} \mathbf{f}).$$

La noción de continuidad puede extenderse de un modo natural a las funciones vectoriales de un vector:

2.4 Definición. *La función* $\mathbf{f}$ *es* ***continua en el punto*** $\mathbf{a}$ *de* $\mathscr{D}_\mathbf{f}$ *si para todo* $\varepsilon > 0$ *existe un* $\delta > 0$ *tal que*

$$|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{a})| < \varepsilon$$

siempre que $\mathbf{x}\in\mathscr{D}_\mathbf{f}$ *y* $|\mathbf{x}-\mathbf{a}| < \delta$.

Si $\mathbf{a}$ no es un punto de acumulación de $\mathscr{D}_\mathbf{f}$, entonces $\mathbf{f}$ es continua en $\mathbf{a}$. Si $\mathbf{a}$ es un punto de acumulación de $\mathscr{D}_\mathbf{f}$, entonces la definición 2.4 es equivalente a la función $\mathbf{f}$ es continua en $\mathbf{a}$ si $\lim_{\mathbf{a}} \mathbf{f} = \mathbf{f}(\mathbf{a})$.

Los siguientes teoremas se siguen fácilmente de los teoremas 2.2 y 2.3.

2.5 Teorema. *La función* $\mathbf{f}$ *es continua en* $\mathbf{a}$ *si y sólo si cada una de sus funciones componentes es continua en* $\mathbf{a}$.

2.6 Teorema. *Si las funciones* $\mathbf{f}$, $\mathbf{g}$ *y* φ *son continuas en* $\mathbf{a}$, *entonces* $\mathbf{f}+\mathbf{g}$, $\mathbf{f}-\mathbf{g}$, $\mathbf{f}\cdot\mathbf{g}$, $\mathbf{f}\times\mathbf{g}$ *y* $\varphi\mathbf{f}$ *son continuas en* $\mathbf{a}$.

Decimos que $\mathbf{f}$ es continua si es continua en cada punto de su dominio y decimos que $\mathbf{f}$ es continua sobre un conjunto $\mathscr{S} \subset \mathscr{D}_\mathbf{f}$ si la función restringida $\mathbf{f}_\mathscr{S}$ es continua. Recordemos que $\mathbf{f}_\mathscr{S}$ es la función con dominio $\mathscr{D}_\mathbf{f} \cap \mathscr{S}$ tal que $\mathbf{f}_\mathscr{S}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ si $\mathbf{x}\in\mathscr{D}_\mathbf{f} \cap \mathscr{S}$.

El siguiente teorema es una generalización del teorema del valor intermedio (teorema 5.7, pág. 187). Probamos primero un lema.

2.7 Lema. *Sea* $\mathbf{f}$ *una función continua de* R^n *en* R^m *con dominio* $\mathscr{D}$. *Si* $\mathscr{A}$ *es un abierto relativo a* $\mathscr{R} = \mathbf{f}(\mathscr{D})$, *entonces* $\mathbf{f}^*(\mathscr{A}) = \{\mathbf{x} \mid \mathbf{f}(\mathbf{x})\in\mathscr{A}\}$ *es abierto respecto a* $\mathscr{D}$.

PRUEBA. Tomemos $\mathbf{x}_0\in\mathbf{f}^*(\mathscr{A})$ y sea $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$. Como $\mathscr{A}$ es abierto relativo a $\mathscr{R}$ y $\mathbf{y}_0\in\mathscr{A}$, existe una vecindad $\mathscr{S}(\mathbf{y}_0;\varepsilon)$ tal que $\mathscr{S}(\mathbf{y}_0;\varepsilon) \cap \mathscr{R} \subset \mathscr{A}$. Por otra parte, como $\mathbf{f}$ es continua en $\mathbf{x}_0$, correspondiéndose con el ε hay una $\delta > 0$ tal que $\mathbf{x}\in\mathscr{S}(\mathbf{x}_0;\delta) \cap \mathscr{D}$ implica $\mathbf{f}(\mathbf{x})\in\mathscr{S}(\mathbf{y}_0;\varepsilon) \cap \mathscr{R} \subset \mathscr{A}$. Así pues $\mathscr{S}(\mathbf{x}_0;\delta) \cap \mathscr{D}$ está contenida en $\mathbf{f}^*(\mathscr{A})$, de donde $\mathbf{f}^*(\mathscr{A})$ es abierto relativo a $\mathscr{D}$.

2.8 Teorema. *Sea* $\mathbf{f}$ *una función continua de* R^n *en* R^m. *Si* $\mathscr{E}$ *es cualquier subconjunto conexo del dominio de* $\mathbf{f}$, *entonces* $\mathbf{f}(\mathscr{E})$ *es conexo.*

PRUEBA. Podemos suponer que $\mathscr{E}$ es el dominio de $\mathbf{f}$. Supongamos ahora que $\mathbf{f}(\mathscr{E})$ no es conexo. Entonces existen dos conjuntos no vacíos ajenos $\mathscr{A}$ y $\mathscr{B}$ ambos abiertos relativamente a $\mathbf{f}(\mathscr{E})$ tales que $\mathbf{f}(\mathscr{E}) = \mathscr{A} \cup \mathscr{B}$. De acuerdo con el lema 2.7 los conjuntos $\mathbf{f}^*(\mathscr{A})$ y $\mathbf{f}^*(\mathscr{B})$ son abiertos relativos a $\mathscr{E}$. Además, estos dos conjuntos son ajenos y no vacíos y

$$\mathscr{E} = \mathbf{f}^*(\mathscr{A} \cup \mathscr{B}) = \mathbf{f}^*(\mathscr{A}) \cup \mathbf{f}^*(\mathscr{B}).$$

Esto significa que $\mathscr{E}$ no es conexo. Esta contradicción implica que $\mathbf{f}(\mathscr{E})$ es conexo.

Problemas

1. Pruébese el teorema 2.2

2. Determínense los siguientes límites:

a) $\lim\limits_{(x,y)\to(1,2)} (x^2 y, x+y, 2x)$

b) $\lim_{(-1,5,2)} (I_1 I_2 I_3, I_1 - I_3)$

c) $\lim_{(x,y)\to(3,2)} \left(\text{sen}\, xy, \tan\frac{y}{x}\right).$

3. Pruébese el teorema 2.3.

4. Si $\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{b}$, pruébese que

$$\lim_{\mathbf{x}\to\mathbf{a}} |\mathbf{f}(\mathbf{x})| = |\mathbf{b}|$$

y

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{f}(\mathbf{x})}{|\mathbf{f}(\mathbf{x})|} = \frac{\mathbf{b}}{|\mathbf{b}|} \quad \text{si } \mathbf{b} \neq \mathbf{0}.$$

5. Si $\mathbf{f}$ tiene la propiedad de que $|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{y})| \leqslant |\mathbf{x}-\mathbf{y}|$ para todo $\mathbf{x}, \mathbf{y} \in \mathscr{D}_{\mathbf{f}}$, demuéstrese que $\mathbf{f}$ es una función continua.

6. Sea $\mathbf{f}$ una función de R^n en R^m con dominio $\mathscr{D}$ y sea $\mathscr{R} = \mathbf{f}(\mathscr{D})$. Si para todo conjunto $\mathscr{A}$ abierto relativamente a $\mathscr{R}$, $\mathbf{f}^*(\mathscr{A})$ es abierto relativamente a $\mathscr{D}$, demuéstrese que $\mathbf{f}$ es continua.

7. Usando el hecho de que un intervalo en R es conexo, demuéstrese que un segmento rectilíneo en R^n es conexo.

8. Pruébese que un conjunto convexo (problema 8, pág. 195) en R^n es conexo.

9. Demuéstrese que una curva punteada, pág. 110, en R^n es conexa.

10. Si $\mathscr{E}$ es un conjunto en R^n con la propiedad de que cualesquier dos puntos de $\mathscr{E}$ pueden unirse por una curva punteada contenida en $\mathscr{E}$, demuéstrese que $\mathscr{E}$ es conexo.

3. MATRICES

Antes de proceder a una discusión de la derivada de una función de R^n en R^m, introduciremos brevemente la noción de matriz de números reales. *Una* ***matriz*** *$m \times n$ de números reales, A, es una función con dominio el conjunto de pares de enteros $\{(i,j) \mid 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n\}$ y con rango en* R.

Un valor de la función $A(i,j)$ se llama *entrada* de la matriz; la entrada $A(i,j)$ suele también denotarse por a_{ij}. Usualmente una matriz se

describe desplegando las entradas en una disposición rectangular; por ejemplo

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Nótese que el arreglo con que describimos la matriz $m \times nA$ tiene m renglones y n columnas; a_{ij} es la entrada en el i-ésimo renglón y j-ésima columna.

Como dos funciones son iguales si y sólo si tienen el mismo dominio y la misma regla de correspondencia, vemos que dos matrices A y B son iguales si y sólo si tienen el mismo número de renglones, el mismo número de columnas y, además, las entradas correspondientes son iguales, es decir, $A(i,j) = B(i,j)$.

Pasamos ahora a definir las operaciones de adición de matrices y multiplicación de una matriz por un número real.

3.1 Definición. *Si A y B son matrices $m \times n$, entonces la suma de A y B es la matriz $m \times n$ $A+B$ tal que $(A+B)(i,j) = A(i,j)+B(i,j)$ para $1 \leqslant i \leqslant m$ y $1 \leqslant j \leqslant n$.*

La suma $A+B$ está definida solamente si A y B tienen el mismo número de renglones y el mismo número de columnas; cada entrada de la suma se obtiene sumando las entradas correspondientes de A y B. Expresando las matrices por sus arreglos rectangulares, tenemos

$$A+B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix}.$$

3.2 Definición. *Si A es una matriz $m \times n$ y r es un número real, entonces rA es la matriz $m \times n$ tal que $[rA](i,j) = rA(i,j)$ para $1 \leqslant i \leqslant m$ y $1 \leqslant j \leqslant n$*

Así pues

$$rA = \begin{pmatrix} ra_{11} & ra_{12} & \dots & ra_{1n} \\ ra_{21} & ra_{22} & \dots & ra_{2n} \\ \vdots & \vdots & & \vdots \\ ra_{m1} & ra_{m2} & \dots & ra_{mn} \end{pmatrix}.$$

Es fácil ver que, para m y n fijos, el conjunto de todas las matrices $m \times n$ de números reales con estas operaciones de adición y multiplicación por un real forman un espacio vectorial sobre el campo real (problema 2). La matriz $m \times n$, O, es la matriz todas cuyas entradas son iguales a cero y $-A$, la inversa aditiva de A, es la matriz con entradas $[-A](i,j) = -A(i,j)$.

Ahora demostraremos que el espacio vectorial constituido por todas las matrices $1 \times n$ de números reales es isomorfo a R^n; es decir, que existe una correspondencia biyectiva (uno a uno y sobre) entre las matrices $1 \times n$ y los vectores de R^n tal que las operaciones de adición y multiplicación por un número real se preservan bajo esta correspondencia. Sea $\mathbf{a} = (a_1, \ldots, a_n)$ el vector correspondiente a $A = (a_{11} \ldots a_{1n})$ si y sólo si $a_j = a_{1j}$ para $1 \leqslant j \leqslant n$. Entonces, si $\mathbf{a}$ y $\mathbf{b}$ se corresponden con A y B respectivamente

$\mathbf{a}+\mathbf{b} = (a_1+b_1, \ldots, a_n+b_n)$ se corresponde con

$$A+B = (a_{11}+b_{11}, \ldots, a_{1n}+b_{1n})$$

y

$r\mathbf{a} = (ra_1, \ldots, ra_n)$ se corresponde con $rA = (ra_{11} \ldots ra_{1n})$.

Como las matrices $1 \times n$ tiene las características de los vectores en R^n identificaremos una matriz $1 \times n$ con el vector correspondiente; las matrices $1 \times n$ se llaman *vectores renglón*. De un modo análogo podemos establecer un isomorfismo entre el espacio de las matrices $n \times 1$ y R^n. Por tanto, una matriz $n \times 1$ puede identificarse con el vector correspondiente en R^n; a las matrices $n \times 1$ se les llama *vectores columna*. En general, el espacio vectorial consistente en todas las matrices $m \times n$ de números reales es isomorfo a R^{mn} (problema 6).

Las definiciones dadas para la adición de matrices y la multiplicación de una matriz por un número real son conformes al modo habitual de sumar y multiplicar por un número real las funciones reales de un vector. Sin embargo, la definición de multiplicación de matrices que ahora daremos difiere esencialmente de la dada para la multiplicación de funciones reales de un vector.

3.3 Definición. *Si A es una matriz $m \times n$ y B es una matriz $n \times p$, el* ***producto de*** *A* ***y*** *B es la matriz $m \times p$ C tal que $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$, para $1 \leqslant i \leqslant m$ y $1 \leqslant j \leqslant p$.*

El producto AB está definido solamente si el número de columnas de A es igual al número de renglones de B. Si convenimos en denotar por $\mathbf{a}_i$ el i-ésimo renglón de A y por $\mathbf{b}^j$ la j-ésima columna de B, entonces podemos considerar $\mathbf{a}_i$ *y* $\mathbf{b}^j$ como vectores de R^n y $c_{ij} = \mathbf{a}_i \cdot \mathbf{b}^j$. Así pues, podemos escribir

3.4
$$AB = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}^1 & \mathbf{a}_1 \cdot \mathbf{b}^2 & \dots & \mathbf{a}_1 \cdot \mathbf{b}^p \\ \mathbf{a}_2 \cdot \mathbf{b}^1 & \mathbf{a}_2 \cdot \mathbf{b}^2 & \dots & \mathbf{a}_2 \cdot \mathbf{b}^p \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_m \cdot \mathbf{b}^1 & \mathbf{a}_m \cdot \mathbf{b}^2 & \dots & \mathbf{a}_m \cdot \mathbf{b}^p \end{pmatrix}.$$

Por ejemplo,
$$\begin{pmatrix} 2 & 5 \\ -1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} -3 & 1 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 14 & 12 \\ 3 & -1 \\ -5 & 5 \end{pmatrix}.$$

Si A es una matriz 1×1 podemos establecer la siguiente correspondencia biunívoca: A se corresponde con su única entrada a_{11}. Es fácil ver que las operaciones de adición y multiplicación se preservan en esta correspondencia entre matrices 1×1 y números reales. Podemos, pues, identificar una matriz 1×1 con el número real que es su única entrada.

Observemos ahora que si A es una matriz $1 \times n$ y B una matriz $n \times 1$ y $\mathbf{a}$ y $\mathbf{b}$ son los vectores correspondientes en R^n, entonces el producto escalar $\mathbf{a} \cdot \mathbf{b}$ se corresponde con AB. Por ejemplo,

$$(2, -1, 5) \cdot (-3, 0, 7) = 29$$

se corresponde con

$$(2 \quad -1 \quad 5) \begin{pmatrix} -3 \\ 0 \\ 7 \end{pmatrix} = (29).$$

Como un tratamiento completo de la multiplicación de matrices nos llevaría mucho tiempo, limitaremos nuestra discusión a las propiedades que vamos a necesitar en nuestro estudio de las derivadas.

3.5 Teorema. *Si A es una matriz $m \times n$, B una matriz $n \times p$, y C es una matriz $p \times q$, entonces $A(BC) = (AB)C$.*

PRUEBA. Tanto $A(BC)$ como $(AB)C$ son matrices $m \times q$. Además, para $1 \leqslant i \leqslant m$ y $1 \leqslant j \leqslant q$,

$$[A(BC)](i, j) = \sum_{k=1}^{n} A(i, k)\,[BC](k, j) = \sum_{k=1}^{n} a_{ik} \sum_{l=1}^{p} b_{kl} c_{lj}$$
$$= \sum_{k=1}^{n} \sum_{l=1}^{p} a_{ik} b_{kl} c_{lj} = \sum_{l=1}^{p} \sum_{k=1}^{n} a_{ik} b_{kl} c_{lj}$$

$$= \sum_{l=1}^{p} \left(\sum_{k=1}^{n} a_{ik} b_{kl} \right) c_{lj} = \sum_{l=1}^{p} [AB](i,l)\, C(l,j)$$

$$= [(AB)C]\,(i,j).$$

Por tanto, $A(BC) = (AB)C$.

Como un resultado de esta ley asociativa de la multiplicación de matrices, podemos escribir el producto de las matrices A, B, y C en la forma ABC, sin que sean necesarios paréntesis algunos para indicar el orden en el que los productos han de efectuarse.

3.6 Teorema. *Si A es una matriz $m \times n$ y B y C son matrices $n \times p$, entonces $A(B+C) = AB+AC$.*

PRUEBA. $A(B+C)$ y $AB+AC$ son matrices $m \times p$. Para todo (i,j) tal que $1 \leqslant i \leqslant m$ y $1 \leqslant j \leqslant p$, tenemos

$$[A(B+C)](i,j) = \sum_{k=1}^{n} A(i,k)\,[B+C](k,j) = \sum_{k=1}^{n} a_{ik}(b_{kj}+c_{kj})$$

$$= \sum_{k=1}^{n} (a_{ik}b_{kj}+a_{ik}c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}$$

$$= [AB+AC]\,(i,j).$$

El teorema 3.6 es una ley distributiva para las matrices. Como el producto de matrices no es conmutativo, deberíamos también probar la siguiente ley distributiva: si A y B son matrices $m \times n$ y C es una matriz $n \times p$, entonces $(A+B)C = AC+BC$. Sin embargo, no desarrollamos la prueba por ser completamente análoga a la del teorema 3.6. Usando estas leyes distributivas obtenemos: si A y B son matrices $m \times n$ y C y D son matrices $n \times p$, entonces $(A+B)(C+D) = AC+BC+AD+BD$.

Introducimos a continuación el concepto de norma de una matriz.

3.7 Definición. *Una función real definida sobre el conjunto $\mathscr{A}$ de todas las matrices con entradas reales, denotada por $\|\ \|$, se llama* ***norma matricial*** *si para todas las matrices A, $B \in \mathscr{A}$ y para todos los números reales r, tenemos*

1) $\|A\| > 0$ *si* $A \neq O$, *y* $\|O\| = 0$
2) $\|rA\| = |r|\, \|A\|$
3) $\|A+B\| \leqslant \|A\| + \|B\|$ *(A y B son matrices $m \times n$)*
4) $\|AB\| \leqslant \|A\|\, \|B\|$ *(A una matriz $m \times n$ y B una matriz $n \times p$).*

Es fácil ver que si identificamos los vectores en el espacio real

n-dimensional con las matrices $n \times 1$ (o con las matrices $1 \times n$), entonces la distancia euclidiana es una norma matricial:

$$|\mathbf{x}| > 0 \text{ si } \mathbf{x} \neq \mathbf{0}, \text{ y } |\mathbf{0}| = 0$$

$$|r\mathbf{x}| = |r|\,|\mathbf{x}|$$

$$|\mathbf{x}+\mathbf{y}| \leqslant |\mathbf{x}|+|\mathbf{y}| \quad \text{(desigualdad del triángulo)}$$

$$|\mathbf{x}\cdot\mathbf{y}| \leqslant |\mathbf{x}|\,|\mathbf{y}| \quad \text{(desigualdad de Schwarz).}$$

Como otro ejemplo podemos observar que para vectores en el espacio real n-dimensional podemos definir una norma vectorial (matricial) por la regla

$$\|\mathbf{x}\| = |x_1|+|x_2|+ \ldots +|x_n| .$$

Que ésta satisface las propiedades (1)-(4) de la definición 3.7, puede verificarse fácilmente. Por ejemplo

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\| &= |x_1+y_1|+|x_2+y_2|+ \ldots +|x_n+y_n| \\ &\leqslant |x_1|+|y_1|+|x_2|+|y_2|+ \ldots +|x_n|+|y_n| = \|\mathbf{x}\|+\|\mathbf{y}\| . \end{aligned}$$

3.8 Teorema. *La función real definida sobre el conjunto $\mathscr{A}$ de todas las matrices de entradas reales por la regla*

$$\|A\| = \left[\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right]^{1/2}$$

donde A es una matriz $m \times n$, es una norma matricial. La llamaremos ***norma matricial euclidiana***

PRUEBA. Si identificamos $A \in \mathscr{A}$, matriz $m \times n$, con un vector en R^{mn} (problema 6), entonces $\|A\|$ es exactamente la longitud euclidiana de este vector, y las propiedades (1), (2) y (3) son las propiedades fundamentales de la longitud de un vector (teorema 5.2, pág. 26). Para probar la propiedad (4), observamos que si $C = AB$, entonces

$$\|A\|^2\,\|B\|^2 = \left[\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right]\left[\sum_{k=1}^{n}\sum_{l=1}^{p} b_{kl}^2\right] = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{p} a_{ij}^2 b_{kl}^2$$

y

$$\begin{aligned} \|AB\|^2 = \|C\|^2 = \sum_{i=1}^{m}\sum_{l=1}^{p} c_{il}^2 &= \sum_{i=1}^{m}\sum_{l=1}^{p}\left[\sum_{j=1}^{n} a_{ij}b_{jl}\right]^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{p} a_{ij}b_{jl}a_{ik}b_{kl} \end{aligned}$$

de modo que

$$\begin{aligned} \|A\|^2\,\|B\|^2 - \|AB\|^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{p} a_{ij}^2 b_{kl}^2 - \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{p} a_{ij}b_{jl}a_{ik}b_{kl} \\ &= \tfrac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{p} (a_{ij}b_{kl}-a_{ik}b_{jl})^2 \geqslant 0 . \end{aligned}$$

De donde

$$\|A\|^2 \, \|B\|^2 \geqslant \|AB\|^2$$

lo que implica la propiedad (4).

En lo que falta de este libro siempre que aparezca la norma de una matriz ha de entenderse que esta es la norma matricial euclidiana del teorema 3.8.

Finalmente, introducimos la noción de límite para funciones matriciales (es decir, con un conjunto de matrices como rango) de un vector. Sea F la función matricial definida por

$$F(\mathbf{y}) = \begin{pmatrix} f_{11}(\mathbf{y}) & \cdots & f_{1n}(\mathbf{y}) \\ \dots & \dots & \dots \\ f_{m1}(\mathbf{y}) & \cdots & f_{mn}(\mathbf{y}) \end{pmatrix}$$

donde cada función f_{ij} es una función real de un vector. Como es usual, en la definición de $\lim_{\mathbf{x}} F$ que sigue suponemos que $\mathbf{x}$ es un punto de acumulación del dominio de F.[1]

3.9 Definición. *Se dice que la matriz A es el* ***límite de la función matricial*** *F* ***en*** *$\mathbf{x}$, lo que se escribe,* $\lim_{\mathbf{x}} F = A$ *o* $\lim_{\mathbf{y}\to\mathbf{x}} F(\mathbf{y}) = A$, *si para cada número $\varepsilon > 0$, hay un número $\delta > 0$, tal que siempre que $\mathbf{y} \in \mathscr{D}_F$ y $0 < |\mathbf{y}-\mathbf{x}| < \delta$ entonces*

$$\|F(\mathbf{y}) - A\| < \varepsilon .$$

Nótese que si F es una función cuyos valores son matrices $m \times n$, entonces A es una matriz $m \times n$.

3.10 Teorema. *Sea A una matriz $m \times n$, F una función cuyos valores son matrices $m \times n$ y dominio un conjunto de vectores, y $\mathbf{x}$ un punto de acumulación de $\mathscr{D}_F$. Entonces* $\lim_{\mathbf{x}} F = A$ *si y sólo si* $\lim_{\mathbf{x}} f_{ij} = a_{ij}$ *para todo $i = 1, \dots, m$ y $j = 1, \dots, n$.*

La prueba se omite ya que es la misma que la prueba del teorema correspondiente para funciones de R en R^n (pág. 103).

Problemas

1. Si

$$A = \begin{pmatrix} -1 & 0 & 1 \\ 3 & 5 & 3 \\ 2 & 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 7 & 1 & 4 \\ 3 & 5 & -2 \\ -2 & 1 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 1 & 5 \\ 2 & 7 & -4 \\ 0 & 2 & 3 \end{pmatrix}$$

[1] En esto y en lo que sigue siempre supondremos que todos los vectores que constituyen el dominio de una función son del mismo espacio vectorial. [N. del T.]

determínense

a) $B+C$ *b*) AB *c*) BA
d) $(AB)C$ *e*) $AB+AC$ *f*) $A(B+C)$.

2. Demuéstrese que el conjunto M de todas las matrices $m \times n$ de números reales es un espacio vectorial sobre el campo real, pág. 37; es decir, demuéstrese que

A_2. $A+B = B+A$ para toda A y B en M.
A_3. $(A+B)+C = A+(B+C)$ para toda A, B y C en M.
A_4. Existe una matriz O en M tal que, para toda A en M, $A+O = A$.
A_5. Para cada A en M existe una matriz $-A$ en M tal que $A+(-A) = O$.
S_2. $1 \cdot A = A$ para toda A en M.
S_3. $r(sA) = (rs)A$ para toda A en M y r y s en R.
S_4. $(r+s)A = rA+sA$ para toda A en M y r y s en R.
S_5. $r(A+B) = rA+rB$ para toda A y B en M y r en R.

3 Si A es una matriz $m \times n$ e I es la matriz $n \times n$ con entradas $I(i,j) = \delta_{ij}$, donde δ_{ij} (llamada delta de Kronecker) es 1 si $i = j$ y es cero si $i \neq j$, demuéstrese que $AI = A$. I se llama la matriz *identidad* $n \times n$.

4. Si

$$A = \begin{pmatrix} 5 & -3 & 4 \\ 1 & 0 & 2 \end{pmatrix} \quad \text{y} \quad \mathbf{b} = \begin{pmatrix} 7 \\ 2 \\ 9 \end{pmatrix}$$

determínese $A\mathbf{b}$.

5. Si r es un número real, A es una matriz $m \times n$, y B es una matriz $n \times p$, demuéstrese que

$$r(AB) = (rA)B = A(rB).$$

6. Demuéstrese que la correspondencia uno-uno

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \leftrightarrow \mathbf{a} = (a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{m1}, \ldots, a_{mn})$$

establece un isomorfismo entre el espacio vectorial M de todas las matrices $m \times n$ de números reales y el espacio vectorial R^{mn}; es decir, demuéstrese que si $A \leftrightarrow \mathbf{a}$ y $B \leftrightarrow \mathbf{b}$, entonces $A+B \leftrightarrow \mathbf{a}+\mathbf{b}$ y $rA \leftrightarrow r\mathbf{a}$ para todo número real r.

7. Demuéstrese que $A\mathbf{x} = B\mathbf{x}$ para toda $\mathbf{x}$ implica $A = B$.

8. Para vectores en R^n definimos, para cada $p \geq 1$, una norma por la regla

$$\|\mathbf{x}\|_p = [|x_1|^p + \ldots + |x_n|^p]^{1/p}$$

y definimos la norma $\|\mathbf{x}\|_\infty$ por la regla

$$\|\mathbf{x}\|_\infty = \text{máx}\,\{|x_i| \mid i = 1, \ldots, n\}.$$

Demuéstrese que

$$\|\mathbf{x}\|_\infty \leqslant \|\mathbf{x}\|_1 \leqslant n\|\mathbf{x}\|_\infty$$

y

$$\|\mathbf{x}\|_\infty \leqslant \|\mathbf{x}\|_2 \leqslant \sqrt{n}\,\|\mathbf{x}\|_\infty .$$

a) Si

$$F(x) = \begin{pmatrix} 3x & x^2 - 1 & 4 \\ x + 1 & x^3 & x^2 \\ 2x^2 & x - 4 & x \end{pmatrix}$$

determínese $\lim_{x \to 2} F(x)$.

4. LA DIFERENCIAL Y LA DERIVADA

4.1 Definición. *La función* $\mathbf{f}$ *de* R^n *en* R^m *es* ***diferenciable en el punto*** $\mathbf{x}$ *si* $\mathbf{f}$ *está definida en una vecindad* $\mathscr{S}(\mathbf{x} : r)$ *de* $\mathbf{x}$ *y, además, existe una matriz* A *(independiente de* $\mathbf{h}$*) tal que para cualquier punto* $\mathbf{x} + \mathbf{h}$ *en* $\mathscr{S}(\mathbf{x} : r)$

4.2 $$\mathbf{f}(\mathbf{x} + \mathbf{h}) = \mathbf{f}(\mathbf{x}) + A\mathbf{h} + \Phi(\mathbf{x} : \mathbf{h})\mathbf{h}$$

donde $\lim_{\mathbf{h} \to 0} \Phi(\mathbf{x} : \mathbf{h}) = O$. *El término* $A\mathbf{h}$ *se llama* ***diferencial de*** $\mathbf{f}$ ***en*** $\mathbf{x}$ *y* $\mathbf{h}$ *y se denota por* $d\mathbf{f}(\mathbf{x} : \mathbf{h})$. *La matriz* A *se llama* ***derivada de*** $\mathbf{f}$ ***en*** $\mathbf{x}$ *y se denota por* $D\mathbf{f}(\mathbf{x})$.

En 4.2 consideramos todos los vectores como vectores columna; A y $\Phi(\mathbf{x} : \mathbf{h})$ son matrices $m \times n$; y O es la matriz cero $m \times n$. Entonces, usando la notación estándar, podemos escribir la ecuación 4.2 como sigue

$$\begin{pmatrix} f_1(\mathbf{x} + \mathbf{h}) \\ \vdots \\ f_m(\mathbf{x} + \mathbf{h}) \end{pmatrix} = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} + \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$

$$+ \begin{pmatrix} \varphi_{11}(\mathbf{x} : \mathbf{h}) & \ldots & \varphi_{1n}(\mathbf{x} : \mathbf{h}) \\ \vdots & & \vdots \\ \varphi_{m1}(\mathbf{x} : \mathbf{h}) & \ldots & \varphi_{mn}(\mathbf{x} : \mathbf{h}) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$

$$= \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} + \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{h} \\ \vdots \\ \mathbf{a}_m \cdot \mathbf{h} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varphi}_1(\mathbf{x};\mathbf{h})\cdot\mathbf{h} \\ \vdots \\ \boldsymbol{\varphi}_m(\mathbf{x};\mathbf{h})\cdot\mathbf{h} \end{pmatrix}$$

$$= \begin{pmatrix} f_1(\mathbf{x}) + \mathbf{a}_1 \cdot \mathbf{h} + \boldsymbol{\varphi}_1(\mathbf{x};\mathbf{h})\cdot\mathbf{h} \\ \vdots \\ f_m(\mathbf{x}) + \mathbf{a}_m \cdot \mathbf{h} + \boldsymbol{\varphi}_m(\mathbf{x};\mathbf{h})\cdot\mathbf{h} \end{pmatrix}.$$

Así pues, la ecuación 4.2 es equivalente a

4.3 $$f_i(\mathbf{x}+\mathbf{h}) = f_i(\mathbf{x}) + \mathbf{a}_i \cdot \mathbf{h} + \boldsymbol{\varphi}_i(\mathbf{x};\mathbf{h}) \cdot \mathbf{h}, \quad 1 \leqslant i \leqslant m,$$

donde

$$\mathbf{a}_i = (a_{i1} \ldots a_{in}) \quad \text{y} \quad \boldsymbol{\varphi}_i(\mathbf{x};\mathbf{h}) = (\varphi_{i1}(\mathbf{x};\mathbf{h}) \ldots \varphi_{in}(\mathbf{x};\mathbf{h})).$$

Si $\lim_{\mathbf{h}\to\mathbf{0}} \Phi(\mathbf{x};\mathbf{h}) = O$ entonces, para todo i $(1 \leqslant i \leqslant m)$, $\lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\varphi}_i(\mathbf{x};\mathbf{h}) = \mathbf{0}$. Así pues, si $\mathbf{f}$ es diferenciable en $\mathbf{x}$, entonces cada una de las funciones componentes f_i es diferenciable en $\mathbf{x}$. Análogamente, si $\lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\varphi}_i(\mathbf{x};\mathbf{h}) = \mathbf{0}$ para todo i entre 1 y m, entonces $\lim_{\mathbf{h}\to\mathbf{0}} \Phi(\mathbf{x};\mathbf{h}) = O$. Esto muestra que $\mathbf{f}$ es diferenciable en $\mathbf{x}$ si cada una de las funciones componentes es diferenciable en $\mathbf{x}$. A continuación enunciamos estos resultados como un teorema.

4.4 Teorema. *La función* $\mathbf{f} = (f_1, \ldots, f_m)$ *de* R^n *en* R^m *es diferenciable en* $\mathbf{x}$ *si y sólo si cada una de las funciones componentes* f_i $(i = 1, \ldots, m)$ *es diferenciable en* $\mathbf{x}$.

Si $\mathbf{f}$ es diferenciable en $\mathbf{x}$ entonces cada una de las funciones componentes f_i es diferenciable en $\mathbf{x}$ y el vector $\mathbf{a}_i$ en 4.3 es $\mathbf{D}f_i(\mathbf{x})$. De donde,

4.5 $$D\mathbf{f}(\mathbf{x}) = \begin{pmatrix} D_1 f_1(\mathbf{x}) & \ldots & D_n f_1(\mathbf{x}) \\ \vdots & & \vdots \\ D_1 f_m(\mathbf{x}) & \ldots & D_n f_m(\mathbf{x}) \end{pmatrix}$$

y

$$d\mathbf{f}(\mathbf{x};\mathbf{h}) = D\mathbf{f}(\mathbf{x})\mathbf{h} = \begin{pmatrix} \mathbf{D}f_1(\mathbf{x})\cdot\mathbf{h} \\ \vdots \\ \mathbf{D}f_m(\mathbf{x})\cdot\mathbf{h} \end{pmatrix} = \begin{pmatrix} df_1(\mathbf{x};\mathbf{h}) \\ \vdots \\ df_m(\mathbf{x};\mathbf{h}) \end{pmatrix}.$$

La función matricial definida por

$$\begin{pmatrix} D_1 f_1 & \ldots & D_n f_1 \\ \vdots & & \vdots \\ D_1 f_m & \ldots & D_n f_m \end{pmatrix}$$

se llama *matriz jacobiana* de la función $\mathbf{f}$ de R^n en R^m. Hemos, pues,

demostrado que si $\mathbf{f}$ es diferenciable en $\mathbf{x}$ entonces la derivada de $\mathbf{f}$ en $\mathbf{x}$ es el valor de la matriz jacobiana de $\mathbf{f}$ en $\mathbf{x}$.

Supongamos que F es una función matricial definida sobre un conjunto abierto $\mathscr{E}$ en R^n. Decimos que F es *continua* en el punto $\mathbf{x}_0$ de $\mathscr{E}$ si $\lim_{\mathbf{x}_0} F = F(\mathbf{x}_0)$. De acuerdo con el teorema 3.10 es fácil ver que F es continua en $\mathbf{x}_0$ si y sólo si cada una de las entradas f_{ij} de F es continua en $\mathbf{x}_0$. Entonces, según teorema 4.4, es fácil deducir el siguiente teorema.

4.6 Teorema. *Sea* $\mathbf{f}$ *una función de* R^n *en* R^m. *Si la matriz jacobiana de* $\mathbf{f}$ *es continua sobre un conjunto abierto* $\mathscr{E}$, *entonces* $\mathbf{f}$ *es diferenciable sobre* $\mathscr{E}$; *es decir, para cada* $\mathbf{x}\in\mathscr{E}$, *hay una vecindad* $\mathscr{S}(\mathbf{x}; r) \subset \mathscr{E}$ *tal que para cualquier* $\mathbf{x}+\mathbf{h}\in\mathscr{S}'(\mathbf{x}; r)$

4.7 $\mathbf{f}(\mathbf{x}+\mathbf{h}) = \mathbf{f}(\mathbf{x})+D\mathbf{f}(\mathbf{x})\mathbf{h}+\Phi(\mathbf{x}; \mathbf{h})\mathbf{h}$, *donde* $\lim_{\mathbf{h}\to\mathbf{0}} \Phi(\mathbf{x}; \mathbf{h}) = O$.

Además, para cualquier conjunto cerrado $\mathscr{F} \subset \mathscr{E}$ *y cualquier* $\varepsilon > 0$ *existe un* $\delta > 0$ *tal que* $\|\Phi(\mathbf{x}; \mathbf{h})\| < \varepsilon$ *siempre que* $\mathbf{x}$, $\mathbf{x}+\mathbf{h}\in\mathscr{F}$ *y* $0 < |\mathbf{h}| < \delta$.

PRUEBA. Tómese $\mathbf{x}\in\mathscr{E}$ y sea $\mathscr{S}(\mathbf{x}; r)$ una vecindad de $\mathbf{x}$ contenida en $\mathscr{E}$. Tómese $\mathbf{h}$ tal que $|\mathbf{h}| < r$. Entonces, usando el teorema del valor medio (teorema 7.8, pág. 200) para $i = 1, \ldots, m$, tenemos

$$\begin{aligned}
f_i(\mathbf{x}+\mathbf{h})-f_i(\mathbf{x}) \\
&= f_i(\mathbf{x}+h_1\mathbf{u}_1+\ldots+h_n\mathbf{u}_n)-f_i(\mathbf{x})\\
&= f_i(\mathbf{x}+h_1\mathbf{u}_1+\ldots+h_n\mathbf{u}_n)-f_i(\mathbf{x}+h_2\mathbf{u}_2+\ldots+h_n\mathbf{u}_n)\\
&\quad + f_i(\mathbf{x}+h_2\mathbf{u}_2+\ldots+h_n\mathbf{u}_n)-f_i(\mathbf{x}+h_3\mathbf{u}_3+\ldots+h_n\mathbf{u}_n)\\
&\quad + \ldots + f_i(\mathbf{x}+h_n\mathbf{u}_n)-f_i(\mathbf{x})\\
&= h_1 D_1 f_i(\mathbf{x}+\theta_{i1}h_1\mathbf{u}_1+h_2\mathbf{u}_2+\ldots+h_n\mathbf{u}_n)\\
&\quad + h_2 D_2 f_i(\mathbf{x}+\theta_{i2}h_2\mathbf{u}_2+h_3\mathbf{u}_3+\ldots+h_n\mathbf{u}_n)\\
&\quad + \ldots + h_n D_n f_i(\mathbf{x}+\theta_{in}h_n\mathbf{u}_n)
\end{aligned}$$

donde $\theta_{ij}\in\langle 0, 1\rangle$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ y $\mathbf{u}_j$ es el vector unitario con 1 como j-ésimo componente y todos los otros componentes 0. Como las derivadas parciales $D_j f_i$ son continuas sobre $\mathscr{E}$,

$$D_j f_i(\mathbf{x}+\theta_{ij}h_j\mathbf{u}_j+\ldots+h_n\mathbf{u}_n) = D_j f_i(\mathbf{x})+\varphi_{ij}(\mathbf{x}; \mathbf{h})$$

donde $\lim_{\mathbf{h}\to\mathbf{0}} \varphi_{ij}(\mathbf{x}; \mathbf{h}) = 0$. Entonces

$$\begin{aligned}
f_i(\mathbf{x}+\mathbf{h})-f_i(\mathbf{x}) &= h_1(D_1 f_i(\mathbf{x})+\varphi_{i1}(\mathbf{x}; \mathbf{h}))+\ldots+h_n(D_n f_i(\mathbf{x})+\varphi_{in}(\mathbf{x}; \mathbf{h}))\\
&= \mathbf{D}f_i(\mathbf{x})\cdot\mathbf{h}+\boldsymbol{\varphi}_i(\mathbf{x}; \mathbf{h})\cdot\mathbf{h}
\end{aligned}$$

y

$$\mathbf{f}(\mathbf{x}+\mathbf{h})-\mathbf{f}(\mathbf{x}) = D\mathbf{f}(\mathbf{x})\mathbf{h}+\Phi(\mathbf{x}; \mathbf{h})\mathbf{h}.$$

Como cada una de las derivadas parciales $D_j f_i$ es continua sobre cualquier conjunto cerrado $\mathscr{F} \subset \mathscr{E}$, es uniformemente continua sobre $\mathscr{F}$ (teorema 7.6,

página 478). De donde para cada $\varepsilon > 0$ existe una $\delta_{ij} > 0$ tal que $\mathbf{x}$, $\mathbf{x}+\mathbf{h} \in \mathscr{F}$ y $|\mathbf{h}| < \delta_{ij}$ implican

$$|D_j f_i(\mathbf{x}+\theta_{ij} h_j \mathbf{u}_j + \ldots + h_n \mathbf{u}_n) - D_j f_i(\mathbf{x})| = |\varphi_{ij}(\mathbf{x};\mathbf{h})| < \frac{\varepsilon}{\sqrt{mn}}.$$

Sea $\delta = \min \{\delta_{ij}\}$. Entonces $\mathbf{x}$, $\mathbf{x}+\mathbf{h} \in \mathscr{F}$ y $0 < |\mathbf{h}| < \delta$ implican

$$\|\Phi(\mathbf{x};\mathbf{h})\| = \left[\sum_{i=1}^{m} \sum_{j=1}^{n} \varphi_{ij}^{2}(\mathbf{x};\mathbf{h})\right]^{1/2} < \left[\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\varepsilon^2}{mn}\right]^{1/2} = \varepsilon .$$

Lo que completa la prueba.

4.7 Definición. *Una función* $\mathbf{f}$ *de* R^n *en* R^m *se dice que pertenece a la clase* C^k *en un conjunto abierto* $\mathscr{E}$, *lo que escribiremos* $\mathbf{f} \in C^k$ *sobre* $\mathscr{E}$, *si cada una de las funciones componentes* f_i $(i = 1, \ldots, m)$ *es de clase* C^k *sobre* $\mathscr{E}$; *es decir, si todas las derivadas parciales k-ésimas de* f_i *son continuas sobre* $\mathscr{E}$ *para todo* $i = 1, \ldots, m$.

4.8 Ejemplo. Si $\mathbf{f}(x, y, z) = (x^2 + yz, z \operatorname{sen} xy)$, demuéstrese que $\mathbf{f}$ es diferenciable en cualquier punto (x, y, z) de R^3 y determínense $D\mathbf{f}(x, y, z)$ y

$$d\mathbf{f}((x, y, z);\ (dx, dy, dz)).$$

SOLUCIÓN. El valor de la matriz jacobiana de $\mathbf{f}$ en cualquier punto (x, y, z) es

$$\begin{pmatrix} 2x & z & y \\ yz \cos xy & xz \cos xy & \operatorname{sen} xy \end{pmatrix}.$$

Como la matriz jacobiana es continua en todos los puntos (x, y, z) de R^3, $\mathbf{f}$ es diferenciable en (x, y, z).

$$D\mathbf{f}(x, y, z) = \begin{pmatrix} 2x & z & y \\ yz \cos xy & xz \cos xy & \operatorname{sen} xy \end{pmatrix}.$$

Además,

$$d\mathbf{f}((x, y, z); (dx, dy, dz)) = \begin{pmatrix} 2x\,dx + z\,dy + y\,dz \\ yz \cos xy\,dx + xz \cos xy\,dy + \operatorname{sen} xy\,dz \end{pmatrix}.$$

Nótese que en el ejemplo 4.8 las únicas técnicas que se utilizaron fueron las de la diferenciación parcial.

Ahora consideraremos algunas propiedades de la diferenciabilidad de una función de R^n en R^m.

4.9 Teorema. *Si* $\mathbf{f}$ *es diferenciable en* $\mathbf{x}$, *entonces* $\mathbf{f}$ *es continua en* $\mathbf{x}$.

PRUEBA. Si $\mathbf{f}$ es diferenciable en $\mathbf{x}$, entonces para cualquier $\mathbf{x}+\mathbf{h}$ en alguna vecindad reducida de $\mathbf{x}$

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) = \mathbf{f}(\mathbf{x}) + A\mathbf{h} + \Phi(\mathbf{x}:\mathbf{h})\mathbf{h} \text{ donde } \lim_{\mathbf{h}\to 0} \Phi(\mathbf{x}:\mathbf{h}) = O$$

Por tanto, $\lim_{\mathbf{h}\to 0} \mathbf{f}(\mathbf{x}+\mathbf{h}) = \mathbf{f}(\mathbf{x})$ y, por tanto, $\mathbf{f}$ es continua en $\mathbf{x}$.

4.10 Teorema. *Si* $\mathbf{f}$ *y* $\mathbf{g}$ *son diferenciables en* $\mathbf{x}$, *entonces* $\mathbf{f}+\mathbf{g}$ *es diferenciable en* $\mathbf{x}$ *y*

$$D[\mathbf{f}+\mathbf{g}](\mathbf{x}) = D\mathbf{f}(\mathbf{x}) + D\mathbf{g}(\mathbf{x}).$$
$$d[\mathbf{f}+\mathbf{g}](\mathbf{x}:\mathbf{h}) = d\mathbf{f}(\mathbf{x}:\mathbf{h}) + d\mathbf{g}(\mathbf{x}:\mathbf{h}).$$

PRUEBA. Como $\mathbf{f}$ y $\mathbf{g}$ son diferenciables en $\mathbf{x}$, existe una vecindad $\mathscr{N}(\mathbf{x}:r)$ de $\mathbf{x}$ tal que para cualquier punto $\mathbf{x}+\mathbf{h}\in\mathscr{N}'(\mathbf{x}:r)$

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) = \mathbf{f}(\mathbf{x}) + D\mathbf{f}(\mathbf{x})\mathbf{h} + \Phi(\mathbf{x}:\mathbf{h})\mathbf{h} \text{ donde } \lim_{\mathbf{h}\to 0} \Phi(\mathbf{x}:\mathbf{h}) = O$$

y

$$\mathbf{g}(\mathbf{x}+\mathbf{h}) = \mathbf{g}(\mathbf{x}) + D\mathbf{g}(\mathbf{x})\mathbf{h} + \Psi(\mathbf{x}:\mathbf{h})\mathbf{h} \text{ donde } \lim_{\mathbf{h}\to 0} \Psi(\mathbf{x}:\mathbf{h}) = O.$$

Luego, si $\mathbf{x}+\mathbf{h}\in\mathscr{N}'(\mathbf{x}:r)$,

$$\begin{aligned}[\mathbf{f}+\mathbf{g}](\mathbf{x}+\mathbf{h}) &= \mathbf{f}(\mathbf{x}) + D\mathbf{f}(\mathbf{x})\mathbf{h} + \Phi(\mathbf{x}:\mathbf{h})\mathbf{h} + \mathbf{g}(\mathbf{x}) + D\mathbf{g}(\mathbf{x})\mathbf{h} + \Psi(\mathbf{x}:\mathbf{h})\mathbf{h} \\ &= [\mathbf{f}+\mathbf{g}](\mathbf{x}) + [D\mathbf{f}(\mathbf{x}) + D\mathbf{g}(\mathbf{x})]\mathbf{h} + [\Phi(\mathbf{x}:\mathbf{h}) + \Psi(\mathbf{x}:\mathbf{h})]\mathbf{h}.\end{aligned}$$

Como $\lim_{\mathbf{h}\to 0} [\Phi(\mathbf{x}:\mathbf{h}) + \Psi(\mathbf{x}:\mathbf{h})] = O$, $\mathbf{f}+\mathbf{g}$ es diferenciable en $\mathbf{x}$ y

$$D[\mathbf{f}+\mathbf{g}](\mathbf{x}) = D\mathbf{f}(\mathbf{x}) + D\mathbf{g}(\mathbf{x}).$$
$$d[\mathbf{f}+\mathbf{g}](\mathbf{x}:\mathbf{h}) = d\mathbf{f}(\mathbf{x}:\mathbf{h}) + d\mathbf{g}(\mathbf{x}:\mathbf{h}).$$

El teorema 4.10 y teoremas análogos para las otras operaciones definidas sobre las funciones vectoriales de un vector pueden probarse simplemente aplicando a las funciones componentes la teoría para funciones de R^n en R desarrollada en el capítulo 4. Sin embargo, escogimos probar los teoremas 4.9 y 4.10 en una forma que nos diera alguna indicación de cómo podríamos desarrollar la teoría para funciones de R^n en R^m sin considerar primero los casos especiales de funciones de R en R, de R en R^n, y de R^n en R. Si hubiésemos considerado el caso general primero, no hubiésemos tenido que probar el mismo resultado en tres o cuatro ocasiones — por ejemplo, que la diferenciabilidad implica continuidad.

Como la definición 4.1 se aplica en particular a funciones de R en R, de R en R^n y de R^n en R, debemos demostrar que esta definición es compatible con las definiciones de diferenciabilidad, de diferencial y de derivada dadas previamente para estos casos especiales.

Si f es una función de R en R, entonces la ecuación 4.2 es

$$(f(x+h)) = (f(x)) + (a)(h) + (\varphi(x,h))(h)$$

donde cada término es una matriz 1×1. Identificando cada matriz 1×1 con su única entrada, obtenemos

$$f(x+h) = f(x) + ah + \varphi(x; h)h.$$

Además, $a = \lim_{h \to 0} \dfrac{f(x+h) - f(x)}{h} = Df(x)$. Por tanto, la definición 4.1 es compatible con las definiciones previamente dadas de diferenciabilidad, diferencial y derivada de funciones reales de variable real.

Si $\mathbf{f}$ es una función de R en R^n, entonces la ecuación 4.2 es

$$\mathbf{f}(x+h) = \mathbf{f}(x) + A(h) + \Phi(x; h)(h).$$

Identificando la matriz 1×1 (h) con el número real h, vemos que la derivada A se corresponderá con $\lim_{h \to 0} \dfrac{1}{h}[\mathbf{f}(x+h) - \mathbf{f}(x)]$. Esto muestra la compatibilidad de nuestra discusión de funciones de R^n en R^m con la de funciones de R en R^n.

Finalmente, si f es una función de R^n en R, entonces la ecuación 4.2 se convierte en

$$(f(\mathbf{x}+\mathbf{h})) = (f(\mathbf{x})) + A\mathbf{h} + \Phi(\mathbf{x}; \mathbf{h})\mathbf{h}.$$

Identificando las matrices 1×1 $(f(\mathbf{x}+\mathbf{h}))$ y $(f(\mathbf{x}))$ con $f(\mathbf{x}+\mathbf{h})$ y $f(\mathbf{x})$ y las matrices $1 \times n$ A y $\Phi(\mathbf{x}; \mathbf{h})$ con los vectores correspondientes $\mathbf{a}$ y $\boldsymbol{\varphi}(\mathbf{x}; \mathbf{h})$ de R^n, obtenemos

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \mathbf{a} \cdot \mathbf{h} + \boldsymbol{\varphi}(\mathbf{x}; \mathbf{h}) \cdot \mathbf{h}$$

lo que está de acuerdo con 6.3, pág. 189.

Problemas

1. Demuéstrese que $\mathbf{f}$ es diferenciable en todos los puntos (x, y, z) de $\mathscr{D}_{\mathbf{f}}$ y determínense $D\mathbf{f}(x, y, z)$ y $d\mathbf{f}((x, y, z); (dx, dy, dz))$ cuando

a) $\mathbf{f}(x, y, z) = (xz, y+z)$

b) $\mathbf{f}(x, y, z) = (xy, y^2, x^2 z)$

c) $\mathbf{f}(x, y, z) = \left(\dfrac{x}{y}, 2y+1, xz^2\right)$

d) $\mathbf{f}(x, y, z) = (x^2 y, ze^x, x+z)$

2. Si $\mathbf{c}$ es una función constante de R^n en R^m, demuéstrese que $D\mathbf{c} = O$

3. Si $\mathbf{f}$ es la función de R^n en R^n con regla de correspondencia $\mathbf{f}(\mathbf{x}) = a\mathbf{x}$, demuéstrese que $D\mathbf{f}(\mathbf{x}) = aI$ donde I es la matriz identidad $n \times n$ ($I(i, j) = 0$ si $i \neq j$ e $I(i, j) = 1$ si $i = j$).

4. Si $\mathbf{f}$ y $\mathbf{g}$ son funciones de R^n en R^m que son diferenciables en $\mathbf{x}$ y u es una función de R^n en R que es diferenciable en $\mathbf{x}$, demuéstrese que

a) $\mathbf{f}\cdot\mathbf{g}$ es diferenciable en $\mathbf{x}$ y

$$d[\mathbf{f}\cdot\mathbf{g}](\mathbf{x};\mathbf{h}) = \mathbf{f}(\mathbf{x})\cdot d\mathbf{g}(\mathbf{x};\mathbf{h})+\mathbf{g}(\mathbf{x})\cdot d\mathbf{f}(\mathbf{x};\mathbf{h}).$$

b) Si $m = 3$, $\mathbf{f}\times\mathbf{g}$ es diferenciable en $\mathbf{x}$ y

$$d[\mathbf{f}\times\mathbf{g}](\mathbf{x};\mathbf{h}) = \mathbf{f}(\mathbf{x})\times d\mathbf{g}(\mathbf{x};\mathbf{h})+d\mathbf{f}(\mathbf{x};\mathbf{h})\times\mathbf{g}(\mathbf{x}).$$

c) $u\mathbf{f}$ es diferenciable en $\mathbf{x}$ y

$$d[u\mathbf{f}](\mathbf{x};\mathbf{h}) = u(\mathbf{x})\, d\mathbf{f}(\mathbf{x};\mathbf{h})+du(\mathbf{x};\mathbf{h})\,\mathbf{f}(\mathbf{x}).$$

5. REGLA DE LA CADENA

Ahora consideraremos la composición de funciones vectoriales de un vector.

5.1 Definición. *Si* $\mathbf{g}$ *es una función de* R^n *en* R^m *y* $\mathbf{f}$ *es una función de* R^m *en* R^p, *entonces la* ***composición de f con g***, *que escribimos:* $\mathbf{f}\circ\mathbf{g}$, *es la función de* R^n *en* R^p *con regla de correspondencia* $(\mathbf{f}\circ\mathbf{g})(\mathbf{x}) = \mathbf{f}(\mathbf{g}(\mathbf{x}))$ *y dominio*

$$\mathscr{D}_{\mathbf{f}\circ\mathbf{g}} = \{\mathbf{x} \mid \mathbf{x}\in\mathscr{D}_{\mathbf{g}},\ \mathbf{g}(\mathbf{x})\in\mathscr{D}_{\mathbf{f}}\}.$$

Si $\mathbf{f} = (f_1,\ldots,f_p)$, entonces $\mathbf{f}\circ\mathbf{g} = (f_1\circ\mathbf{g},\ldots,f_p\circ\mathbf{g})$.

Antes de considerar la regla de la cadena para diferenciar una composición de funciones, damos algunos teoremas relativos a los límites y a la continuidad de funciones compuestas.

5.2 Teorema. *Si* $\mathbf{g}$ *es una función de* R^n *en* R^m *tal que* $\lim_{\mathbf{a}} \mathbf{g} = \mathbf{b}$, $\mathbf{f}$ *es una función de* R^m *en* R^p *que es continua en* $\mathbf{b}$, *y* $\mathbf{a}$ *es un punto de acumulación del dominio de* $\mathbf{f}\circ\mathbf{g}$, *entonces* $\lim_{\mathbf{a}}(\mathbf{f}\circ\mathbf{g}) = \mathbf{f}(\mathbf{b})$.

5.3 Teorema. *Si* $\mathbf{g}$ *es una función de* R^n *en* R^m *que es continua en* $\mathbf{a}$ *y* $\mathbf{f}$ *es una función de* R^m *en* R^p *que es continua en* $\mathbf{g}(\mathbf{a})$, *entonces* $\mathbf{f}\circ\mathbf{g}$ *es continua en* $\mathbf{a}$.

Las pruebas de estos teoremas son las mismas que las de los teoremas 4.6, pág. 176, y 5.3, pág. 186.

Enunciamos ahora la regla de la cadena para la diferenciación de las funciones compuestas.

5.4 Teorema. *Sea* $\mathbf{g}$ *una función de* R^n *en* R^m *que es diferenciable en un conjunto abierto* $\mathscr{S}$, *y sea* $\mathbf{f}$ *una función de* R^m *en* R^p *que es diferenciable sobre*

un conjunto abierto que contiene $\mathbf{g}(\mathscr{E})$. *Entonces* $\mathbf{f}\circ\mathbf{g}$ *es diferenciable sobre* $\mathscr{E}$ *y para toda* $\mathbf{x}\in\mathscr{E}$ *se verifican las siguientes fórmulas:*

5.5 $$D[\mathbf{f}\circ\mathbf{g}](\mathbf{x}) = D\mathbf{f}(\mathbf{g}(\mathbf{x}))\,D\mathbf{g}(\mathbf{x})$$

y

$$d[\mathbf{f}\circ\mathbf{g}](\mathbf{x};\mathbf{h}) = D\mathbf{f}(\mathbf{g}(\mathbf{x}))\,d\mathbf{g}(\mathbf{x};\mathbf{h}) = d\mathbf{f}(\mathbf{g}(\mathbf{x}); d\mathbf{g}(\mathbf{x};\mathbf{h})).$$

PRUEBA. Sea $\mathbf{x}\in\mathscr{E}$. Como f es diferenciable en $\mathbf{g}(\mathbf{x})$, existe una $p\times m$ matriz A tal que para todos los puntos $\mathbf{g}(\mathbf{x})+\mathbf{k}$ en cierta vecindad reducida $\mathscr{S}'(\mathbf{g}(\mathbf{x}); s)$ de $\mathbf{g}(\mathbf{x})$,

5.6 $$\mathbf{f}(\mathbf{g}(\mathbf{x})+\mathbf{k}) = \mathbf{f}(\mathbf{g}(\mathbf{x}))+[A+\Phi(\mathbf{k})]\mathbf{k} \quad \text{donde} \lim_{\mathbf{k}\to\mathbf{0}} \Phi(\mathbf{k}) = O.$$

Definimos ahora $\Phi(\mathbf{0})$ como la $p\times m$ matriz cero O y observamos que Φ es entonces continua en $\mathbf{0}$. Nótese también que 5.6 se verificará entonces para todo $\mathbf{g}(\mathbf{x})+\mathbf{k}\in\mathscr{S}(\mathbf{g}(\mathbf{x}); s)$. Como $\mathbf{g}$ es diferenciable y, por tanto, continua en $\mathbf{x}$ existe una vecindad $\mathscr{S}(\mathbf{x}; r)$ de $\mathbf{x}$ tal que $\mathbf{g}(\mathscr{S}(\mathbf{x}; r))\subset\mathscr{S}(\mathbf{g}(\mathbf{x}); s)$ y existe una matriz $m\times n$ B tal que, para todo $\mathbf{x}+\mathbf{h}\in\mathscr{S}'(\mathbf{x}; r)$,

5.7 $$\mathbf{g}(\mathbf{x}+\mathbf{h}) = \mathbf{g}(\mathbf{x})+[B+\Psi(\mathbf{h})]\mathbf{h} \quad \text{donde} \lim_{\mathbf{h}\to\mathbf{0}} \Psi(\mathbf{h}) = O.$$

Tomemos ahora $\mathbf{x}+\mathbf{h}$ en $\mathscr{S}'(\mathbf{x}; r)$ y sea $\mathbf{k}(\mathbf{h}) = \mathbf{g}(\mathbf{x}+\mathbf{h})-\mathbf{g}(\mathbf{x})$. Entonces $\lim_{\mathbf{h}\to\mathbf{0}} \mathbf{k}(\mathbf{h}) = \mathbf{0}$. Y de 5.6 y 5.7 obtenemos

$$\begin{aligned}[\mathbf{f}\circ\mathbf{g}](\mathbf{x}+\mathbf{h}) &= \mathbf{f}(\mathbf{g}(\mathbf{x})+\mathbf{k}(\mathbf{h}))\\ &= \mathbf{f}(\mathbf{g}(\mathbf{x}))+[A+\Phi(\mathbf{k}(\mathbf{h}))]\mathbf{k}(\mathbf{h})\\ &= \mathbf{f}(\mathbf{g}(\mathbf{x}))+[A+\Phi(\mathbf{k}(\mathbf{h}))]\,[B+\Psi(\mathbf{h})]\mathbf{h}\\ &= \mathbf{f}(\mathbf{g}(\mathbf{x}))+AB\mathbf{h}+\Theta(\mathbf{h})\mathbf{h}.\end{aligned}$$

Como

$$\lim_{\mathbf{h}\to\mathbf{0}} \Theta(\mathbf{h}) = \lim_{\mathbf{h}\to\mathbf{0}} [\Phi(\mathbf{k}(\mathbf{h}))B+A\Psi(\mathbf{h})+\Phi(\mathbf{k}(\mathbf{h}))\Psi(\mathbf{h})] = O,$$

$\mathbf{f}\circ\mathbf{g}$ es diferenciable en $\mathbf{x}$. Usando el hecho de que $A = D\mathbf{f}(\mathbf{g}(\mathbf{x}))$ y $B = D\mathbf{g}(\mathbf{x})$, obtenemos

$$D[\mathbf{f}\circ\mathbf{g}](\mathbf{x}) = D\mathbf{f}(\mathbf{g}(\mathbf{x}))D\mathbf{g}(\mathbf{x}).$$

Además

$$\begin{aligned}d[\mathbf{f}\circ\mathbf{g}](\mathbf{x};\mathbf{h}) &= D\mathbf{f}(\mathbf{g}(\mathbf{x}))D\mathbf{g}(\mathbf{x})\mathbf{h} = D\mathbf{f}(\mathbf{g}(\mathbf{x}))\,d\mathbf{g}(\mathbf{x};\mathbf{h})\\ &= d\mathbf{f}(\mathbf{g}(\mathbf{x}); d\mathbf{g}(\mathbf{x};\mathbf{h})).\end{aligned}$$

Lo que completa la prueba.

Por 5.5 vemos que la entrada $D_j[f_i\circ\mathbf{g}](\mathbf{x})$ en el i-ésimo renglón y j-ésima columna de $D(\mathbf{f}\circ\mathbf{g})(\mathbf{x})$ es el i-ésimo renglón de $D\mathbf{f}(\mathbf{g}(\mathbf{x}))$ por la j-ésima columna de $D\mathbf{g}(\mathbf{x})$, es decir,

5.8 $$D_j[f_i\circ\mathbf{g}](\mathbf{x}) = \mathbf{D}f_i(\mathbf{g}(\mathbf{x}))\cdot D_j\mathbf{g}(\mathbf{x})$$

donde $D_j\mathbf{g} = (D_jg_1, \ldots, D_jg_m)$. Llamamos también a 5.8 la regla de la cadena ya que es equivalente a 5.5

Desarrollaremos ahora un caso particular de la regla de la cadena 5.5 en la notación de variables. Supongamos que $\mathbf{g}$ es una función de R^2 en R^3 y que f es una función de R^3 en R. Sea $F = f \circ \mathbf{g}$, $(x, y, z) = \mathbf{g}(u, v)$, y $w = F(u, v) = f(x, y, z)$. Entonces, usando la regla de la cadena tenemos

$$\begin{pmatrix} \dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} & \dfrac{\partial f}{\partial z} \end{pmatrix} \begin{pmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \\ \dfrac{\partial z}{\partial u} & \dfrac{\partial z}{\partial v} \end{pmatrix}$$

es decir,

$$\frac{\partial F}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial u} \quad \text{y} \quad \frac{\partial F}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial v}$$

o bien

$$\frac{\partial w}{\partial u} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial u} \quad \text{y} \quad \frac{\partial w}{\partial v} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial v}.$$

5.9 Ejemplo. Sea f una función real que es diferenciable sobre R^2. Si describimos los puntos de R^2 por medio de coordenadas polares, entonces f se transforma en la función $F = f \circ \mathbf{g}$ donde $\mathbf{g}(r, \theta) = (r\cos\theta, r\operatorname{sen}\theta)$; es decir,

$$F(r, \theta) = f(r\cos\theta, r\operatorname{sen}\theta).$$

Encuéntrense expresiones para las derivadas parciales de F en términos de las derivadas parciales de f.

SOLUCIÓN 1. Usando la regla de la cadena, tenemos

$$\mathbf{D}F(r, \theta) = \mathbf{D}f(r\cos\theta, r\operatorname{sen}\theta)\,D\mathbf{g}(r, \theta)$$

$$= (D_1 f(r\cos\theta, r\operatorname{sen}\theta), D_2 f(r\cos\theta, r\operatorname{sen}\theta)) \begin{pmatrix} D_1 g_1(r, \theta) & D_2 g_1(r, \theta) \\ D_1 g_2(r, \theta) & D_2 g_2(r, \theta) \end{pmatrix}$$

$$= (D_1 f(r\cos\theta, r\operatorname{sen}\theta), D_2 f(r\cos\theta, r\operatorname{sen}\theta)) \begin{pmatrix} \cos\theta & -r\operatorname{sen}\theta \\ \operatorname{sen}\theta & r\cos\theta \end{pmatrix}$$

de modo que

$$D_1 F(r, \theta) = \cos\theta\, D_1 f(r\cos\theta, r\operatorname{sen}\theta) + \operatorname{sen}\theta\, D_2 f(r\cos\theta, r\operatorname{sen}\theta)$$

y

$$D_2 F(r, \theta) = -r \operatorname{sen} \theta\, D_1 f(r \cos \theta, r \operatorname{sen} \theta) + r \cos \theta\, D_2 f(r \cos \theta, r \operatorname{sen} \theta).$$

Si hacemos $(x, y) = \mathbf{g}(r, \theta)$, entonces $F(r, \theta) = f(x, y)$, y la solución puede escribirse como sigue:

$$\mathbf{D}F(r, \theta) = \left(\frac{\partial F}{\partial r}, \frac{\partial F}{\partial \theta}\right) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \begin{pmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{pmatrix} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \begin{pmatrix} \cos \theta & -r \operatorname{sen} \theta \\ \operatorname{sen} \theta & r \cos \theta \end{pmatrix}$$

de modo que

$$\frac{\partial F}{\partial r} = \cos \theta \frac{\partial f}{\partial x} + \operatorname{sen} \theta \frac{\partial f}{\partial y} \quad \text{y} \quad \frac{\partial F}{\partial \theta} = -r \operatorname{sen} \theta \frac{\partial f}{\partial x} + r \cos \theta \frac{\partial f}{\partial y}.$$

Solución 2. Podemos obtener estos resultados usando diferenciales. Por definición

$$dF((r, \theta); (dr, d\theta)) = D_1 F(r, \theta) dr + D_2 F(r, \theta) d\theta.$$

Usando la regla de la cadena tenemos

$$\begin{aligned} dF((r, \theta); (dr, d\theta)) &= df(\mathbf{g}(r, \theta); d\mathbf{g}((r, \theta); (dr, d\theta))) \\ &= D_1 f(\mathbf{g}(r, \theta)) dg_1((r, \theta); (dr, d\theta)) \\ &\quad + D_2 f(\mathbf{g}(r, \theta)) dg_2((r, \theta); (dr, d\theta)) \\ &= (\cos \theta\, dr - r \operatorname{sen} \theta\, d\theta)\, D_1 f(r \cos \theta, r \operatorname{sen} \theta) \\ &\quad + (\operatorname{sen} \theta\, dr + r \cos \theta\, d\theta)\, D_2 f(r \cos \theta, r \operatorname{sen} \theta) \\ &= [\cos \theta\, D_1 f(r \cos \theta, r \operatorname{sen} \theta) + \operatorname{sen} \theta\, D_2 f(r \cos \theta, r \operatorname{sen} \theta)] dr \\ &\quad + [-r \operatorname{sen} \theta\, D_1 f(r \cos \theta, r \operatorname{sen} \theta) + r \cos \theta\, D_2 f(r \cos \theta, \\ &\qquad r \operatorname{sen} \theta)] d\theta. \end{aligned}$$

Por tanto,

$$D_1 F(r, \theta) = \cos \theta\, D_1 f(r \cos \theta, r \operatorname{sen} \theta) + \operatorname{sen} \theta\, D_2 f(r \cos \theta, r \operatorname{sen} \theta),$$
$$D_2 F(r, \theta) = -r \operatorname{sen} \theta\, D_1 f(r \cos \theta, r \operatorname{sen} \theta) + r \cos \theta\, D_2 f(r \cos \theta, r \operatorname{sen} \theta).$$

Usando la notación de variables, podemos escribir la segunda solución del ejemplo 5.9 como sigue. Sea $(x, y) = \mathbf{g}(r, \theta) = (r \cos \theta, r \operatorname{sen} \theta)$ y $z = F(r, \theta) = f(x, y)$. Entonces

$$dz = \frac{\partial F}{\partial r} dr + \frac{\partial F}{\partial \theta} d\theta.$$

Además,

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial}{\partial y}dy$$

$$= (\cos\theta\, dr - r \operatorname{sen}\theta\, d\theta)\frac{\partial f}{\partial x} + (\operatorname{sen}\theta\, dr + r\cos\theta\, d\theta)\frac{\partial f}{\partial y}$$

$$= \left(\cos\theta\frac{\partial f}{\partial x} + \operatorname{sen}\theta\frac{\partial f}{\partial y}\right)dr + \left(-r\operatorname{sen}\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}\right)d\theta.$$

Por tanto,

$$\frac{\partial F}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \operatorname{sen}\theta\frac{\partial f}{\partial y},$$

$$\frac{\partial F}{\partial \theta} = -r\operatorname{sen}\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}.$$

Podemos prescindir por completo de los símbolos de función y escribir:

$$dz = \frac{\partial z}{\partial r}dr + \frac{\partial z}{\partial \theta}d\theta$$

y

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy = \frac{\partial z}{\partial x}(\cos\theta\, dr - r\operatorname{sen}\theta\, d\theta) + \frac{\partial z}{\partial y}(\operatorname{sen}\theta\, dr + r\cos\theta\, d\theta)$$

$$= \left(\cos\theta\frac{\partial z}{\partial x} + \operatorname{sen}\theta\frac{\partial z}{\partial y}\right)dr + \left(-r\operatorname{sen}\theta\frac{\partial z}{\partial x} + r\cos\theta\frac{\partial z}{\partial y}\right)d\theta.$$

El resultado se escribiría entonces:

$$\frac{\partial z}{\partial r} = \cos\theta\frac{\partial z}{\partial x} + \operatorname{sen}\theta\frac{\partial z}{\partial y},$$

$$\frac{\partial z}{\partial \theta} = -r\operatorname{sen}\theta\frac{\partial z}{\partial x} + r\cos\theta\frac{\partial z}{\partial y}.$$

Ilustramos ahora el uso de la regla de la cadena para la determinación de las derivadas parciales de orden mayor.

5.10 Ejemplo. Si $F = f \circ \mathbf{g}$ donde $\mathbf{g}(r, \theta) = (r\cos\theta, r\operatorname{sen}\theta)$, proporciónese una expresión para $D_{2,1}F$ en términos de las derivadas parciales de f. Se supone que las derivadas parciales de segundo orden de f son continuas.

SOLUCIÓN. Como las derivadas parciales de segundo orden de f son continuas, las derivadas parciales de primer orden de f y la propia f son diferenciables. En la solución del ejemplo 5.9 demostramos que

$$D_1 F(r, \theta) = \cos\theta\, D_1 f(r\cos\theta, r\operatorname{sen}\theta) + \operatorname{sen}\theta\, D_2 f(r\cos\theta, r\operatorname{sen}\theta).$$

Luego,

$$\begin{aligned}
&D_{2,1} F(r, \theta)\\
&\quad = \cos\theta[\mathbf{D} D_1 f(r\cos\theta, r\operatorname{sen}\theta)\cdot(-r\operatorname{sen}\theta, r\cos\theta)]\\
&\qquad - \operatorname{sen}\theta\, D_1 f(r\cos\theta, r\operatorname{sen}\theta)\\
&\qquad + \operatorname{sen}\theta[\mathbf{D} D_2 f(r\cos\theta, r\operatorname{sen}\theta)\cdot(-r\operatorname{sen}\theta, r\cos\theta)]\\
&\qquad + \cos\theta\, D_2 f(r\cos\theta, r\operatorname{sen}\theta)\\
&\quad = -r\operatorname{sen}\theta\cos\theta\, D_{1,1} f(r\cos\theta, r\operatorname{sen}\theta) + r\cos^2\theta\, D_{2,1} f(r\cos\theta, r\operatorname{sen}\theta)\\
&\qquad - \operatorname{sen}\theta\, D_1 f(r\cos\theta, r\operatorname{sen}\theta) - r\operatorname{sen}^2\theta\, D_{1,2} f(r\cos\theta, r\operatorname{sen}\theta)\\
&\qquad + r\operatorname{sen}\theta\cos\theta\, D_{2,2} f(r\cos\theta, r\operatorname{sen}\theta) + \cos\theta\, D_2 f(r\cos\theta, r\operatorname{sen}\theta),
\end{aligned}$$

Como f tiene derivadas parciales de segundo orden continuas,

$$D_{1,2} f(r\cos\theta, r\operatorname{sen}\theta) = D_{2,1} f(r\cos\theta, r\operatorname{sen}\theta)$$

de donde,

$$\begin{aligned}
D_{2,1} F(r, \theta) &= r\operatorname{sen}\theta\cos\theta[D_{2,2} f(r\cos\theta, r\operatorname{sen}\theta) - D_{1,1} f(r\cos\theta, r\operatorname{sen}\theta)]\\
&\quad + r(\cos^2\theta - \operatorname{sen}^2\theta) D_{2,1} f(r\cos\theta, r\operatorname{sen}\theta) - \operatorname{sen}\theta\, D_1 f(r\cos\theta, r\operatorname{sen}\theta)\\
&\quad + \cos\theta\, D_2 f(r\cos\theta, r\operatorname{sen}\theta).
\end{aligned}$$

Desarrollando la solución en la notación de variables, tenemos

$$\frac{\partial F}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \operatorname{sen}\theta\frac{\partial f}{\partial y}$$

y

$$\begin{aligned}
\frac{\partial^2 F}{\partial\theta\,\partial r} &= \frac{\partial}{\partial\theta}\left(\frac{\partial F}{\partial r}\right) = \cos\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) - \operatorname{sen}\theta\frac{\partial f}{\partial x} + \operatorname{sen}\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right) + \cos\theta\frac{\partial f}{\partial y}\\
&= \cos\theta\left[\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial\theta} + \frac{\partial^2 f}{\partial y\,\partial x}\frac{\partial y}{\partial\theta}\right] - \operatorname{sen}\theta\frac{\partial f}{\partial x}\\
&\qquad + \operatorname{sen}\theta\left[\frac{\partial^2 f}{\partial x\,\partial y}\frac{\partial x}{\partial\theta} + \frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial\theta}\right] + \cos\theta\frac{\partial f}{\partial y}\\
&= -r\operatorname{sen}\theta\cos\theta\frac{\partial^2 f}{\partial x^2} + r\cos^2\theta\frac{\partial^2 f}{\partial y\,\partial x} - \operatorname{sen}\theta\frac{\partial f}{\partial x} - r\operatorname{sen}^2\theta\frac{\partial^2 f}{\partial x\,\partial y}
\end{aligned}$$

$$+ r \operatorname{sen}\theta \cos\theta \frac{\partial^2 f}{\partial y^2} + \cos\theta \frac{\partial f}{\partial y}$$

$$= r \operatorname{sen}\theta \cos\theta \left(\frac{\partial^2 f}{\partial y^2} - \frac{\partial^2 f}{\partial x^2}\right) + r(\cos^2\theta - \operatorname{sen}^2\theta) \frac{\partial^2 f}{\partial x \partial y}$$

$$- \operatorname{sen}\theta \frac{\partial f}{\partial x} + \cos\theta \frac{\partial f}{\partial y}.$$

Pasamos ahora a otro tipo de problemas. Deseamos demostrar que, bajo ciertas condiciones, las ecuaciones

5.11 $$F(x, y, u, v) = 0 \quad \text{y} \quad G(x, y, u, v) = 0$$

pueden resolverse para dos de las variables en términos de las otras dos, digamos $u = f(x, y)$ y $v = g(x, y)$. Obtendremos también expresiones para las derivadas parciales de f y g en términos de las derivadas parciales de F y G.

Supongamos que F y G son funciones de R^4 en R que pertenecen a la clase C^1 en un conjunto abierto $\mathscr{E}$ y que para algún punto $\mathbf{P}_0 = (x_0, y_0, u_0, v_0)$ en $\mathscr{E}$,

$$F(\mathbf{P}_0) = 0, \quad G(\mathbf{P}_0) = 0, \quad \text{y} \quad \begin{vmatrix} D_3F(\mathbf{P}_0) & D_4F(\mathbf{P}_0) \\ D_3G(\mathbf{P}_0) & D_4G(\mathbf{P}_0) \end{vmatrix} \neq 0.$$

Este determinante de derivadas parciales se llama jacobiano y se representa por $\dfrac{\partial(F, G)}{\partial(u, v)}$. Como el jacobiano es distinto de cero en $\mathbf{P}_0$, podemos suponer que $D_3G(\mathbf{P}_0)$ y $D_4F(\mathbf{P}_0)$ son distintas de cero [el procedimiento sería análogo si $D_3F(\mathbf{P}_0)$ y $D_4G(\mathbf{P}_0)$ fueran distintas de cero]. Usando el análogo del teorema 13.3, pág. 229 (el teorema de la función implícita) para funciones de R^4 en R, vemos que existe una vecindad $\mathscr{U}$ de (x_0, y_0, u_0) y una función $h \in C^1$ sobre $\mathscr{U}$ tal que $v_0 = h(x_0, y_0, u_0)$ y $F(x, y, u, h(x, y, u)) = 0$ para todo $(x, y, u) \in \mathscr{U}$. Además, para todo $(x, y, u) \in \mathscr{U}$

5.12 $$D_i h(x, y, u) = -\frac{D_i F(x, y, u, h(x, y, u))}{D_4 F(x, y, u, h(x, y, u))}, \quad i = 1, 2, 3.$$

Si hacemos $H(x, y, u) = G(x, y, u, h(x, y, u))$, entonces para $(x, y, u) \in \mathscr{U}$

$$\begin{aligned} D_i H(x, y, u) &= D_i G(x, y, u, h(x, y, u)) + D_4 G(x, y, u, h(x, y, u)) D_i h(x, y, u) \\ &= D_i G(x, y, u, h(x, y, u)) \\ &\quad - D_4 G(x, y, u, h(x, y, u)) \frac{D_i F(x, y, u, h(x, y, u))}{D_4 F(x, y, u, h(x, y, u))} \end{aligned}$$

y con las variables omitidas por brevedad,

5.13 $$D_i H = \frac{D_i G D_4 F - D_4 G D_i F}{D_4 F}, \quad i = 1, 2, 3.$$

Como $H \in C^1$ sobre $\mathscr{M}$, $H(x_0, y_0, u_0) = 0$, y $D_3 H(x_0, y_0, u_0) \neq 0$, el teorema de la función implícita, pág. 229, implica la existencia de una vecindad $\mathscr{N}$ de (x_0, y_0), una vecindad $\langle u_0 - c, u_0 + c \rangle$ de u_0, y una función $f \in C^1$ sobre $\mathscr{N}$ tal que la región cilíndrica

$$\{(x, y, u) \mid (x, y) \in \mathscr{N}, \ u \in \langle u_0 - c, u_0 + c \rangle\} \subset \mathscr{M},$$
$$u_0 = f(x_0, y_0),$$

y, para todo $(x, y) \in \mathscr{N}$

$$f(x, y) \in \langle u_0 - c, u_0 + c \rangle \ \text{ y } \ H(x, y, f(x, y)) = 0.$$

Además, para toda $(x, y) \in \mathscr{N}$,

5.14 $$D_i f(x, y) = -\frac{D_i H(x, y, f(x, y))}{D_3 H(x, y, f(x, y))}, \quad i = 1, 2.$$

Haciendo $g(x, y) = h(x, y, f(x, y))$, tenemos para toda $(x, y) \in \mathscr{N}$

$$F(x, y, f(x, y), g(x, y)) = 0 \ \text{ y } \ G(x, y, f(x, y), g(x, y)) = 0;$$

es decir, $u = f(x, y)$ y $v = g(x, y)$ son soluciones de 5.11 en alguna vecindad de (x_0, y_0). Por 5.13 y 5.14

5.15 $$D_i f(x, y) = -\frac{D_i G D_4 F - D_4 G D_i F}{D_3 G D_4 F - D_4 G D_3 F}, \quad i = 1, 2$$

donde las funciones a la derecha están evaluadas en $(x, y, f(x, y), g(x, y))$. Además

$$D_i g(x, y) = D_i h(x, y, f(x, y)) + D_3 h(x, y, f(x, y)) D_i f(x, y), \quad i = 1, 2.$$

Usando 5.12 y 5.15, obtenemos

$$D_i g(x, y) = -\frac{D_i F}{D_4 F} + \frac{D_3 F}{D_4 F} \cdot \frac{D_i G D_4 F - D_4 G D_i F}{D_3 G D_4 F - D_4 G D_3 F}$$

5.16

$$= -\frac{D_i F D_3 G - D_3 F D_i G}{D_3 G D_4 F - D_4 G D_3 F} \quad i = 1, 2$$

donde las funciones a la derecha están evaluadas en $(x, y, f(x, y), g(x, y))$.

Hemos probado que si F y G pertenecen a la clase C^1 en un conjunto abierto S y si el jacobiano $\frac{\partial(F,G)}{\partial(u,v)}$ es distinto de cero en S entonces en una vecindad de cada punto de S las ecuaciones

$$F(x,y,u,v)=0 \quad \text{y} \quad G(x,y,u,v)=0$$

pueden resolverse para u y v en términos de x y y y

$$\frac{\partial u}{\partial x} = -\frac{\dfrac{\partial(F,G)}{\partial(x,v)}}{\dfrac{\partial(F,G)}{\partial(u,v)}} \qquad \frac{\partial u}{\partial y} = -\frac{\dfrac{\partial(F,G)}{\partial(y,v)}}{\dfrac{\partial(F,G)}{\partial(y,v)}}$$

$$\frac{\partial v}{\partial x} = -\frac{\dfrac{\partial(F,G)}{\partial(u,x)}}{\dfrac{\partial(F,G)}{\partial(u,v)}} \qquad \frac{\partial v}{\partial y} = -\frac{\dfrac{\partial(F,G)}{\partial(u,y)}}{\dfrac{\partial(F,G)}{\partial(u,v)}}.$$

Resultados análogos son válidos para otros casos. En general, bajo condiciones apropiadas, m ecuaciones en n variables ($n > m$) pueden resolverse para m de las variables en términos de las otras variables. Si suponemos la existencia de una solución tal, podemos calcular las derivadas parciales de un modo formal. Ilustraremos esto en el siguiente ejemplo.

5.17 Ejemplo. Supongamos que las ecuaciones

$$x^2 + yu + 2z - v = 0$$
$$xv + y^2 + 3zu^2 - 5 = 0$$

pueden resolverse para u y v en términos de x, y y z. Determínense $\frac{\partial u}{\partial x}$ y $\frac{\partial v}{\partial x}$.

Solución. Suponiendo que u y v son funciones de x, y, y z, tomamos las derivadas parciales de las ecuaciones con respecto a x:

$$2x + y\frac{\partial u}{\partial x} - 2v\frac{\partial v}{\partial x} = 0$$

$$x\frac{\partial v}{\partial x} + v + 6zu\frac{\partial u}{\partial x} = 0.$$

Resolviendo estas ecuaciones para $\dfrac{\partial u}{\partial x}$ y $\dfrac{\partial v}{\partial x}$, obtenemos

$$\frac{\partial u}{\partial x} = \frac{\begin{vmatrix} -2x & -2v \\ -v & x \end{vmatrix}}{\begin{vmatrix} y & -2v \\ 6zu & x \end{vmatrix}} = -2\,\frac{x^2+v^2}{xy+12zuv}$$

$$\frac{\partial v}{\partial x} = \frac{\begin{vmatrix} y & -2x \\ 6zu & -v \end{vmatrix}}{\begin{vmatrix} y & -2v \\ 6zu & x \end{vmatrix}} = \frac{12xzu-yv}{xy+12zuv}.$$

Problemas

1. Si $(x, y) = \mathbf{g}(u, v, w)$ y $z = f(x, y) = f(\mathbf{g}(u, v, w))$, desarróllense las fórmulas para $\dfrac{\partial z}{\partial u}$, $\dfrac{\partial z}{\partial v}$, y $\dfrac{\partial z}{\partial w}$ de acuerdo con la regla de la cadena.

2. Supongamos $\mathbf{g}(u, v, w) = (uv^2, v+3w^2)$, $f(x, y) = x^2 - y^2$ y $z = f(\mathbf{g}(u, v, w))$. Encuéntrense $\dfrac{\partial z}{\partial u}$, $\dfrac{\partial z}{\partial v}$, y $\dfrac{\partial z}{\partial w}$ por cálculo directo y también aplicando la regla de la cadena.

3. Si $\mathbf{g}(u, v, w) = (u \operatorname{sen} v, w^2, u^2 w)$, $f(x, y, z) = xyz$, y $t = f(\mathbf{g}(u, v, w))$, encuéntrense $\dfrac{\partial t}{\partial u}$, $\dfrac{\partial t}{\partial v}$, y $\dfrac{\partial t}{\partial w}$ por cálculo directo y también por aplicación de la regla de la cadena.

4. Si f es una función diferenciable y $F = f \circ \mathbf{g}$ donde $\mathbf{g}(\rho, \theta, \varphi) = (\rho \operatorname{sen} \varphi \cos \theta, \rho \operatorname{sen} \varphi \operatorname{sen} \theta, \rho \cos \varphi)$, encuéntrese una relación entre las derivadas parciales de F y las de f. (F es la función en que f se transforma bajo el cambio de coordenadas rectangulares a esféricas.)

5. Igual al problema 4, salvo que $\mathbf{g}(r, \theta, z) = (r \cos \theta, r \operatorname{sen} \theta, z)$. (Este problema ocurre cuando cambiamos de coordenadas rectangulares a cilíndricas.)

6. Determínense las derivadas parciales de $F = f \circ \mathbf{g}$ cuando

a) $\mathbf{g}(u, v) = (u \cos a \cos v, u \cos a \operatorname{sen} v, u \operatorname{sen} a)$
$f(x, y, z) = x^2 + y^2$

b) $\mathbf{g}(u, v) = (u \cos v, u \operatorname{sen} v, 3u^2)$
$f(x, y, z) = \dfrac{xy - z}{yz}$.

7. Supongamos que f tiene derivadas parciales continuas de segundo orden. Si $u = f(x, y)$, entonces $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = 0$ se llama ecuación de Laplace en dos dimensiones. Demuéstrese que si cambiamos a coordenadas polares de modo que $u = f(r \cos \theta, r \operatorname{sen} \theta)$, entonces la ecuación de Laplace toma la forma

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0.$$

8. Supongamos que f tiene derivadas parciales continuas de segundo orden y $u = f(x, t)$. Demuéstrese que la ecuación de onda $\dfrac{\partial^2 u}{\partial t^2} - a^2 \dfrac{\partial^2 u}{\partial x^2} = 0$ toma la forma $\dfrac{\partial^2 u}{\partial r \partial s} = 0$ bajo el cambio de variables: $r = x + at$, $s = x - at$.

9. Demuéstrese que $u = F(x - at) + G(x + at)$, donde F y G son funciones arbitrarias con derivadas parciales segundas continuas, satisfacen la ecuación de onda

$$\frac{\partial^2 u}{\partial t^2} - a^2 \frac{\partial^2 u}{\partial x^2} = 0$$

10. Si $u = F(x - y, y - z, z - x)$ demuéstrese que $\dfrac{\partial u}{\partial x} + \dfrac{\partial u}{\partial y} + \dfrac{\partial u}{\partial z} = 0.$

11. Demuéstrese que la ecuación

$$x^2 \frac{\partial^2 u}{\partial x^2} + y^2 \frac{\partial^2 u}{\partial y^2} + x \frac{\partial u}{\partial x} + y \frac{\partial u}{\partial y} + 0$$

toma la forma $\dfrac{\partial^2 u}{\partial r^2} + \dfrac{\partial^2 u}{\partial s^2} = 0$ bajo el cambio de variables: $x = e^r$, $y = e^s$.

12. Si u, v, x y y están relacionadas por las ecuaciones

$$xyu - yv^2 + x^2 = 0 \qquad 4u^2 + 2v^2 - x^3 y = 0$$

encuéntrense:

$a)$ $\dfrac{\partial u}{\partial x}, \dfrac{\partial u}{\partial y}, \dfrac{\partial v}{\partial x}$, y $\dfrac{\partial v}{\partial y}$ $\qquad$ $b)$ $\dfrac{\partial x}{\partial u}, \dfrac{\partial x}{\partial v}, \dfrac{\partial y}{\partial u}$, y $\dfrac{\partial y}{\partial v}$.

13. Si las variables u, v, x, y y z están relacionadas por las ecuaciones

$$x;u;+xzv+y; = 0 \qquad yzu+xyv;-3x = 0$$

encuéntrense $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}, \frac{\partial v}{\partial z}$.

6. SUPERFICIES

En el estudio de las superficies nos enfrentamos con un problema análogo a uno con que nos encontramos al discutir las curvas: una superficie, ¿qué va a ser?, ¿un conjunto de puntos o una función? En conformidad con nuestro tratamiento de las curvas elegimos definir una superficie como una función o, lo que es equivalente, un conjunto de puntos descrito de una forma particular.

6.1 Definición. *Una* ***superficie*** *en* R^n *es una función continua de un subconjunto conexo de* R^2 *en* R^n.

Nosostros aquí vamos a considerar tan solo superficies en R^3 y, por tanto, el término "superficie" significará una función continua de R^2 en R^3.

Asociado a una superficie $\mathbf{f}$ siempre tenemos un conjunto de puntos en R^3: el rango de $\mathbf{f}$. Podemos considerar la superficie como este conjunto de puntos descrito en la forma particular determinada por $\mathbf{f}$. Denotamos por ello a una superficie por $\mathscr{S}$, por ejemplo, y decimos que $\mathscr{S}$ es la superficie descrita por $\mathbf{f}$. Si $\mathscr{S}$ está descrita por $\mathbf{f}$, entonces la ecuación $\mathbf{x} = \mathbf{f}(u, v)$ se llama ecuación paramétrica de $\mathscr{S}$.

Por ejemplo, si $\mathbf{f}$ tiene como dominio

$$\mathscr{E} = \{(u, v) \mid u \in [0, 2\pi],\ \ v \in \langle -\infty, \infty \rangle\}$$

y regla de correspondencia

$$\mathbf{f}(u, v) = (a \cos u,\, a \operatorname{sen} u,\, v), \text{ donde } a > 0,$$

entonces la superficie descrita por la transformación $\mathbf{f}$ de $\mathscr{E}$ es el cilindro con ecuaciones paramétricas

$$x = a \cos u,\ y = a \operatorname{sen} u,\ z = v;\ u \in [0, 2\pi],\ \ v \in \langle -\infty, \infty \rangle.$$

Podemos ver que $\mathscr{S}$ es un cilindro circular de radio a con el eje Z como eje observando que la distancia del eje Z a cualquier punto (x, y, z) en $\mathscr{S}$ es

$$\sqrt{x^2 + y^2} = \sqrt{a^2 \cos^2 u + a^2 \operatorname{sen}^2 u} = a.$$

Bajo la transformación $\mathbf{f}$ de la faja vertical $\mathscr{E}$, cada recta vertical $u = u_0$ en $\mathscr{E}$ se transforma sobre la recta vertical

$$\mathscr{C}_{u_0} = \{(a \cos u_0,\, a \operatorname{sen} u_0,\, v) \mid v \in \langle -\infty, \infty \rangle\}$$

en R^3 (figura 2); éstos son los "generadores" del cilindro. Cada segmento rectilíneo horizontal [illegible] se transforma sobre la circunferencia

$$[illegible] \; 2\pi]$$

en R^3. Vemos con esto cómo el cilindro [illegible] está formado por la transformación f de la faja [illegible] la faja se estira hasta la longitud $2\pi a$ y se enrolla para formar el cilindro; los bordes [illegible] $= 0$ y $u = 2\pi$ se unen.

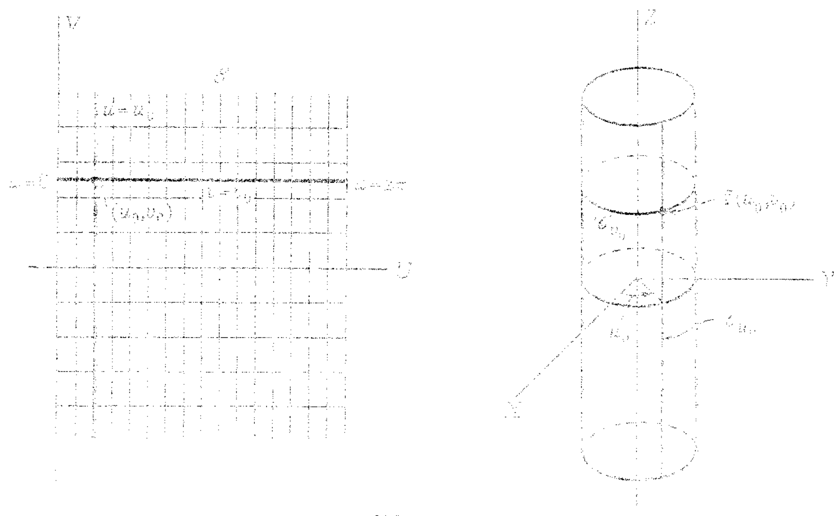


FIGURA 2

[illegible] puntos de la [illegible]

$$[illegible]$$

[illegible] superficie. [illegible] descrito por la transformación [illegible] es la superficie con ecuaciones paramétricas

$$[illegible] = [illegible]$$

Además, si $F \in C^1$ y la derivada de F es distinta de cero en un punto, entonces el teorema de la función implícita (pág. [illegible]) asegura que en una vecindad [illegible] la ecuación [illegible] $= 0$ puede resolverse para una de las variables en términos de las otras dos, digamos [illegible]. Así pues, en una vecindad el conjunto de puntos [illegible] representará la superficie descrita por [illegible].

En la sección 12 del capítulo [illegible] definimos el plano tangente a una superficie [illegible] en el punto [illegible] el plano

que pasa por (x_0, y_0, z_0) con normal $(D_1 g(x_0, y_0), D_2 g(x_0, y_0), -1)$. Haciendo $\mathbf{f} = (I_1, I_2, g)$ vemos que

$$\begin{aligned} D_1 \mathbf{f}(x_0, y_0) \times D_2 \mathbf{f}(x_0, y_0) &= (1, 0, D_1 g(x_0, y_0)) \times (0, 1, D_2 g(x_0, y_0)) \\ &= -(D_1 g(x_0, y_0), D_2 g(x_0, y_0), -1). \end{aligned}$$

Así pues, en este caso $D_1 \mathbf{f}(x_0, y_0) \times D_2 \mathbf{f}(x_0, y_0)$ es una normal al plano tangente en el punto (x_0, y_0, z_0).

Consideremos ahora el caso general de una superficie $\mathscr{S}$ descrita por la transformación $\mathbf{f}$ de $\mathscr{E}$ donde $\mathbf{f} \in C^1$. Sea $\mathscr{C}_{u_0}$ la curva en R^3 con ecuación paramétrica $\mathbf{x} = \mathbf{f}(u_0, v)$, $(u_0, v) \in \mathscr{E}$ y $\mathscr{C}_{v_0}$ la curva con ecuación $\mathbf{x} = \mathbf{f}(u, v_0)$, $(u, v_0) \in \mathscr{E}$ (figura 3); $\mathscr{C}_{u_0}$ suele llamarse curva V-coordenada sobre $\mathscr{S}$ y $\mathscr{C}_{v_0}$ curva U-coordenada sobre $\mathscr{S}$. El vector $D_1 \mathbf{f}(u_0, v_0)$ es un vector tangente a la curva $\mathscr{C}_{v_0}$ en el punto $\mathbf{f}(u_0, v_0)$ y $D_2 \mathbf{f}(u_0, v_0)$ es un vector tangente a $\mathscr{C}_{u_0}$ en $\mathbf{f}(u_0, v_0)$. Si $D_1 \mathbf{f}(u_0, v_0) \times D_2 \mathbf{f}(u_0, v_0) \neq 0$, entonces $D_1 \mathbf{f}(u_0, v_0)$

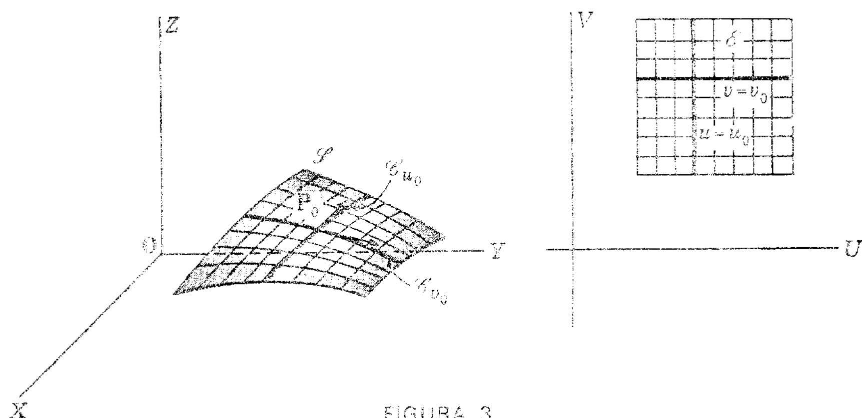


FIGURA 3

y $D_2 \mathbf{f}(u_0, v_0)$ no son paralelos y, por tanto, estos vectores determinan un plano que pasa por $\mathbf{f}(u_0, v_0)$. El plano que pasa por $\mathbf{f}(u_0, v_0)$ con normal $D_1 \mathbf{f}(u_0, v_0) \times D_2 \mathbf{f}(u_0, v_0)$ se llama *plano tangente* de la superficie $\mathscr{S}$ en el punto $\mathbf{f}(u_0, v_0)$. Si $\mathbf{f} \in C^1$ y $D_1 \mathbf{f} \times D_2 \mathbf{f}$ es distinto de cero sobre $\mathscr{E}$, entonces $\mathscr{S}$ tiene un plano tangente en cada punto y se llama *superficie lisa.*

Demostramos que si $\mathscr{L}$ es la recta tangente a una curva lisa situada en la superficie lisa $\mathscr{S}$ que pasa por $\mathbf{x}_0$ entonces $\mathscr{L}$ se encuentra en el plano tangente a $\mathscr{S}$ en $\mathbf{x}_0$. Supongamos a $\mathscr{S}$ descrita por la transformación $\mathbf{f}$ de $\mathscr{E}$ y sea $\mathscr{C}$ una curva en $\mathscr{E}$ descrita por la transformación $\mathbf{g}$ de $\mathscr{I}$ que pasa por el punto (u_0, v_0), digamos que $\mathbf{g}(t_0) = (u_0, v_0)$. Entonces la curva $\mathscr{C}'$ descrita por la transformación $\mathbf{f} \circ \mathbf{g}$ de $\mathscr{I}$ se encuentra en $\mathscr{S}$ y $\mathbf{f}(\mathbf{g}(t_0)) = \mathbf{x}_0$. Por la regla de la cadena

$$D[\mathbf{f} \circ \mathbf{g}](t_0) = D\mathbf{f}(\mathbf{g}(t_0))\, D\mathbf{g}(t_0) = D\mathbf{f}(u_0, v_0)\, D\mathbf{g}(t_0).$$

Omitiendo los puntos en que estas derivadas son evaluadas, tenemos

$$D[\mathbf{f}\circ\mathbf{g}] = \begin{pmatrix} D_1 f_1 & D_2 f_1 \\ D_1 f_2 & D_2 f_2 \\ D_1 f_3 & D_2 f_3 \end{pmatrix} \begin{pmatrix} Dg_1 \\ Dg_2 \end{pmatrix} = \begin{pmatrix} D_1 f_1 Dg_1 + D_2 f_1 Dg_2 \\ D_1 f_2 Dg_1 + D_2 f_2 Dg_2 \\ D_1 f_3 Dg_1 + D_2 f_3 Dg_2 \end{pmatrix}$$

$$= Dg_1 \begin{pmatrix} D_1 f_1 \\ D_1 f_2 \\ D_1 f_3 \end{pmatrix} + Dg_2 \begin{pmatrix} D_2 f_1 \\ D_2 f_2 \\ D_2 f_3 \end{pmatrix}.$$

Así pues

$$D[\mathbf{f}\circ\mathbf{g}](t_0) = Dg_1(t_0)\, D_1\mathbf{f}(u_0, v_0) + Dg_2(t_0)\, D_2\mathbf{f}(u_0, v_0).$$

Esto demuestra que el vector $D(\mathbf{f}\circ\mathbf{g})(t_0)$ que es tangente a la curva $\mathscr{C}'$ en x_0 es una combinación lineal de los vectores $D_1\mathbf{f}(u_0, v_0)$ y $D_2\mathbf{f}(u_0, v_0)$ que determinan el plano tangente a la superficie $\mathscr{S}$ en $\mathbf{x}_0$ y, por tanto, que la recta tangente a $\mathscr{C}'$ en $\mathbf{x}_0$ se encuentra en el plano tangente a $\mathscr{S}$ en $\mathbf{x}_0$.

6.2 Ejemplo. Encuéntrese el plano tangente al cilindro circular de radio 1 con eje el eje Z en el punto $\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 10\right)$.

SOLUCIÓN. El cilindro está descrito por la transformación

$$\mathbf{f}(u, v) = (\cos u, \operatorname{sen} u, v).$$

El punto $\mathbf{x} = \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 10\right)$ se corresponde con $u = \frac{5\pi}{4}$ y $v = 10$. Como $D_1\mathbf{f}(u, v) = (-\operatorname{sen} u, \cos u, 0)$ y $D_2\mathbf{f}(u, v) = (0, 0, 1)$,

$$D_1\mathbf{f}\left(\frac{5\pi}{4}, 10\right) \times D_2\mathbf{f}\left(\frac{5\pi}{4}, 10\right) = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0\right) \times (0,0,1)$$

$$= -\frac{\sqrt{2}}{2}(1,1,0).$$

Así pues, el plano tangente al cilindro en $\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 10\right)$ es el plano que pasa por $\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 10\right)$ con normal $(1,1,0)$. Una ecuación

implícita de este plano es

$$(1,1,0)\cdot\left[(x,y,z)-\left(-\frac{\sqrt{2}}{2},\ -\frac{\sqrt{2}}{2},\ 10\right)\right]=0$$

$$x+y+\sqrt{2}=0\,.$$

Una ecuación vectorial del plano tangente es

$$\mathbf{x}=\left(-\frac{\sqrt{2}}{2},\ -\frac{\sqrt{2}}{2},\ 10\right)+s(1,-1,0)+t(0,0,1)$$

y las correspondientes ecuaciones paramétricas son:

$$x=-\frac{\sqrt{2}}{2}+s,\qquad y=-\frac{\sqrt{2}}{2}-s,\qquad z=10+t\,.$$

Una clase importante de superficies son las superficies de revolución Sea $\mathscr{C}$ una curva en el plano XZ dada por las ecuaciones paramétricas:

$$x=g(v),\ y=0,\ z=h(v),\ \text{ donde }\ v\in\mathscr{J}\,.$$

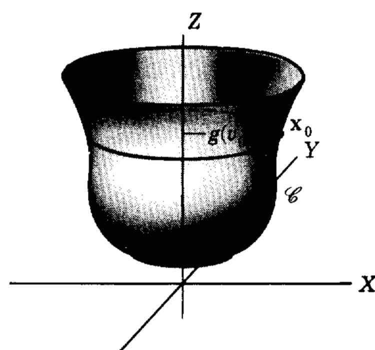


FIGURA 4

Entonces, para cualquier v_0 fijado, la circunferencia obtenida al hacer girar el punto $\mathbf{x}_0=(g(v_0),0,h(v_0))$ alrededor del eje Z (figura 4) tiene ecuaciones paramétricas:

$$x=g(v_0)\cos u,\ y=g(v_0)\operatorname{sen}u,\ z=h(v_0),\ \text{ donde }\ u\in[0,2\pi]\,.$$

El parámetro u es el ángulo de rotación alrededor del eje Z medido desde la dirección positiva del eje X ($u=0$ corresponde al punto $\mathbf{x}_0$). Haciendo girar todos los puntos de $\mathscr{C}$ alrededor del eje Z, obtenemos la superficie

de revolución generada por $\mathscr{C}$. Las ecuaciones paramétricas de esta superficie son:

$$x = g(r)\cos u$$
$$y = g(r)\,\text{sen}\, u$$
$$z = h(r), \quad u \in [0, 2\pi], \quad r \in I$$

Por ejemplo, si $\mathscr{C}$ es la recta vertical con ecuaciones paramétricas

$$x = a, \; y = 0, \; z = r, \quad \text{donde} \quad r \in \langle -\infty, \infty \rangle,$$

la superficie de revolución generada por $\mathscr{C}$ es el *cilindro* con ecuaciones paramétricas

$$x = a\cos u$$
$$y = a\,\text{sen}\, u$$
$$z = r, \quad u \in [0, 2\pi], \quad r \in \langle -\infty, \infty \rangle.$$

Si $\mathscr{C}$ es la recta que pasa por el origen con ecuaciones paramétricas,

$$x = ar, \; y = 0, \; z = r, \quad \text{donde} \quad r \in \langle -\infty, \infty \rangle,$$

entonces la superficie de revolución generada por el giro de $\mathscr{C}$ alrededor del eje Z es el cono de ecuaciones paramétricas:

$$x = ar\cos u$$
$$y = ar\,\text{sen}\, u$$
$$z = r, \quad u \in [0, 2\pi], \quad r \in \langle -\infty, \infty \rangle.$$

Ahora damos una lista de representaciones paramétricas de algunas superficies. Si la superficie tiene una representación sencilla explícita o implícita también la damos.

Elipsoide (figura 5).

$$x = a\cos v\cos u \qquad\qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

$$y = b\cos v\,\text{sen}\, u$$

$$z = c\,\text{sen}\, v, \quad u \in [0, 2\pi], \quad v \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Si $a = b = c$ entonces obtenemos una esfera.

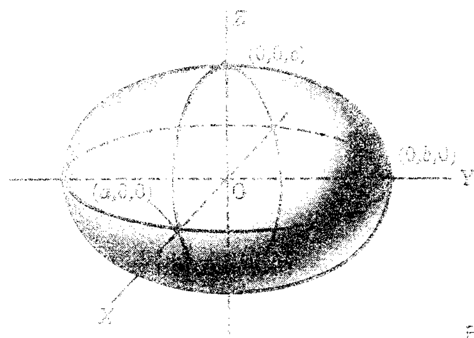


FIGURA 5

Hiperboloide de una hoja (figura 17, pág. 224)

$$x = a \cosh v \cos u \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

$$y = b \cosh v \operatorname{sen} u$$

$$z = c \operatorname{senh} v, \quad u \in [0, 2\pi], \ v \in \langle -\infty, \infty \rangle .$$

Hiperboloide de dos hojas (figura 6)

$$x = a \operatorname{senh} v \cos u \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$

$$y = b \operatorname{senh} v \operatorname{sen} u$$

$$z = c \cosh v, \quad u \in [0, 2\pi], \ v \in \langle -\infty, \infty \rangle .$$

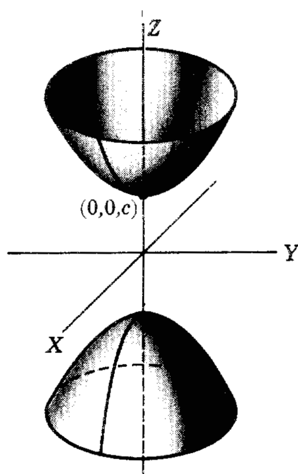


FIGURA 6

Paraboloide elíptico (figura 18, pág. 186)

$$x = av \cos u \qquad z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$

$$y = bv \operatorname{sen} u$$

$$z = v^2, \quad u \in [0, 2\pi], \ v \in [0, \infty \rangle .$$

Si $a = b$ tenemos un *paraboloide circular*.

Paraboloide hiperbólico (figura 24b, pág. 238)

$$x = av \cos u \qquad z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$$

$$y = bv \operatorname{sen} u$$

$$z = v^2 \cos 2u, \quad u \in [0, 2\pi], \ v \in [0, \infty \rangle .$$

Cilindro elíptico recto

$$x = a \cos u \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

$$y = b \operatorname{sen} u$$

$$z = v, \quad u \in [0, 2\pi], \ v \in \langle -\infty, \infty \rangle .$$

Si $a = b$, tenemos un *cilindro circular recto.*

Cono elítico recto

$$x = av \cos u \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} - z^2 = 0$$

$$y = bv \operatorname{sen} u$$

$$z = v, \quad u \in [0, 2\pi], \ v \in \langle -\infty, \infty \rangle .$$

Si $a = b$, tenemos un *cono circular recto.*

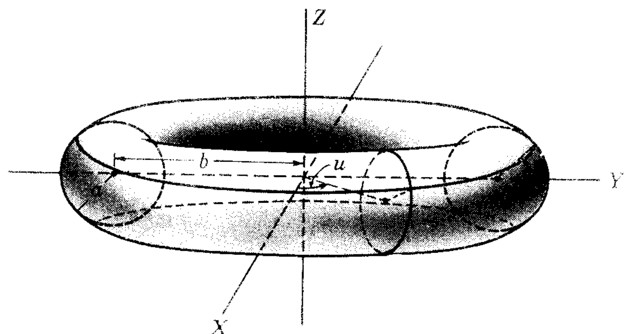


FIGURA 7

Toro (figura 7)

$$x = (b + a \cos v) \cos u$$

$$y = (b + a \cos v) \operatorname{sen} u$$

$$z = a \operatorname{sen} v, \quad u \in [0, 2\pi], \ v \in [0, 2\pi], \ \text{donde} \ a < b.$$

Problemas

1. Considérese el cono circular recto descrito por la transformación $\mathbf{f}$ de $\mathscr{E}$ donde

$$\mathbf{f}(u, v) = (v \cos u, v \operatorname{sen} u, v)$$

y

$$\mathscr{E} = \{(u, v) \mid u \in [0, 2\pi], \ v \in \langle -\infty, \infty \rangle\}$$

a) Determínense las curvas V-coordenadas $\mathscr{C}_{u_0}$ y las curvas U-coordenadas $\mathscr{C}_{v_0}$.
b) Dígase cómo se forma el cono por la transformación $\mathbf{f}$ de $\mathscr{E}$.
c) Encuéntrese el plano tangente del cono en el punto (0, 1, 1).
d) ¿Tiene el cono un plano tangente en (0, 0, 0)?

2. Determínese el plano tangente a la superficie $\mathscr{S}$ en el punto $\mathbf{x}_0$ cuando

a) $\mathscr{S}$ es el cilindro circular de radio 10 alrededor del eje Z; $\mathbf{x}_0 = (5\sqrt{2}, 5\sqrt{2}, 0)$.
b) $\mathscr{S}$ es la esfera de radio $\sqrt{2}$ alrededor del origen; $\mathbf{x}_0 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}, 1\right)$.
c) $\mathscr{S}$ es el paraboloide hiperbólico de ecuaciones

$$\begin{aligned} x &= v \cos u \\ y &= v \operatorname{sen} u \\ z &= v^2 \cos 2u, \quad u \in [0, 2\pi], \; v \in [0, \infty\rangle \end{aligned}$$

y $\mathbf{x}_0 = (\sqrt{2}, \sqrt{2}, 0)$.

d) $\mathscr{S}$ es el elipsoide de ecuaciones

$$\begin{aligned} x &= 2 \cos v \cos u \\ y &= \cos v \operatorname{sen} u \\ z &= \operatorname{sen} v, \quad u \in [0, 2\pi], \; v \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \end{aligned}$$

y $\mathbf{x}_0 = (0, 1, 0)$.

3. Demuéstrese que una esfera es la superficie de revolución generada por la rotación de un círculo alrededor de un diámetro del círculo.

4. Demuéstrese que un toro es la superficie de revolución generada por la rotación de una circunferencia alrededor de una recta que no intersecta a la circunferencia.

5. Determínense ecuaciones paramétricas para las superficies generadas por la rotación de:

a) una elipse alrededor de su eje mayor,
b) una parábola alrededor de su eje,
c) una parábola alrededor de la recta que pasa por su vértice y es paralela a su directriz,
d) una parábola alrededor de su directriz.

6. Determínese la curva que es la intersección de las superficies dadas por las ecuaciones:

$$x^3 - y + 3 = 0$$

y

$$x^3 - z - 2 = 0.$$

7. MULTIPLICADORES DE LAGRANGE

En el capítulo 4 dimos un método para determinar los valores extremos de una función F de R^3 en R sujeta a una restricción: $G(x, y, z) = 0$. Consistía éste en la resolución de la ecuación de restricción para una de la variables, digamos $z = g(x, y)$, para encontrar luego los valores extremos de $f(x, y) = F(x, y, g(x, y))$. Consideraremos ahora otro método para la resolución de tales problemas, debido al matemático francés José Luis Lagrange (1736-1813).

Cuando no hay restricción alguna, sabemos (teorema 4.3, pág. 236) que si una función F de R^n en R definida sobre un conjunto abierto $\mathcal{O}$ tiene un valor extremo en $\mathbf{x}_0 \in \mathcal{O}$ y si $\mathbf{D}F(\mathbf{x}_0)$ existe, entonces $\mathbf{D}F(\mathbf{x}_0) = \mathbf{0}$. Implica esto que si $\mathscr{C}$ es una curva lisa en $\mathcal{O}$ descrita por la función $\mathbf{f}$ donde $\mathbf{f}(t_0) = \mathbf{x}_0$, entonces en $\mathbf{x}_0$ la derivada de F a lo largo de la curva $\mathscr{C}$ es cero; es decir,

$$D[F \circ \mathbf{f}](t_0) = \mathbf{D}F(\mathbf{x}_0) \cdot \mathbf{f}'(t_0) = 0.$$

Consideremos ahora el caso en que deseamos encontrar los valores extremos de F sobre $\mathcal{O}$ sujetos a las restricciones $G_i(\mathbf{x}) = 0$ $(i = 1, \ldots, k)$ donde $\mathbf{D}G_i$ es distinta de cero sobre $\mathcal{O}$. Es decir, queremos determinar los valores extremos de F cuando restringimos F a la intersección de los conjuntos

$$\mathscr{S}_i = \{\mathbf{x} : \mathbf{x} \in \mathcal{O}, G_i(\mathbf{x}) = 0\} \quad (i = 1, \ldots, k).$$

Si $\mathbf{x}_0 \in \mathscr{S}_i$ y la curva lisa $\mathscr{C}$ descrita por $\mathbf{f}$ se encuentra en $\mathscr{S}_i$, entonces $G_i(\mathbf{f}(t)) = 0$ y, por ello,

7.1 $$\mathbf{D}G_i(\mathbf{x}_0) \cdot \mathbf{f}'(t_0) = 0.$$

Así pues, si $\mathbf{D}F(\mathbf{x}_0)$ es una combinación lineal de $\mathbf{D}G_i(\mathbf{x}_0)$ $(i = 1, \ldots, k)$, digamos

$$\mathbf{D}F(\mathbf{x}_0) = \sum_{i=1}^{k} (-\lambda_i)\mathbf{D}G_i(\mathbf{x}_0),$$

entonces 7.1 implica que

$$D[F \circ \mathbf{f}](t_0) = \mathbf{D}F(\mathbf{x}_0) \cdot \mathbf{f}'(t_0) = 0$$

es decir, la derivada de F a lo largo de cualquier curva lisa que se encuentre en $\bigcap_{i=1}^{k} \mathscr{S}_i$ será cero en $\mathbf{x}_0$.

Así pues, desde un punto de vista geométrico esperamos que un valor extremo de F sujeto a las restricciones $G_i(\mathbf{x}) = 0$ sólo puede ocurrir en un punto $\mathbf{x}_0$ tal que

7.2 $$\mathbf{D}F(\mathbf{x}_0) + \sum_{i=1}^{k} \lambda_i \mathbf{D}G_i(\mathbf{x}_0) = \mathbf{0} \quad \text{y} \quad G_i(\mathbf{x}_0) = 0.$$

Si definimos una función H de R^{n+k} en R por

$$H(x_1, ..., x_n, \lambda_1, ..., \lambda_k) = F(x_1, ..., x_n) + \sum_{i=1}^{k} \lambda_i G_i(x_1, ..., x_n)$$

donde $(x_1, \ldots, x_n) \in \mathscr{E}$ y $\lambda_i \in R$, entonces

$$\mathbf{D}H = \left(D_1 F + \sum_{i=1}^{k} \lambda_i D_1 G_i, \ ..., D_n F + \sum_{i=1}^{k} \lambda_i D_n G_i, \ G_1, ..., G_k \right)$$

y, por tanto, las relaciones 7.2 se verifican si y sólo si $\mathbf{D}H(\mathbf{x}_0) = \mathbf{0}$. Así pues, parece que un valor extremo de F sujeto a las restricciones $G_i(\mathbf{x}) = 0$ puede ocurrir solamente en un punto $\mathbf{x} = (x_1, \ldots, x_n)$ solamente si $(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_k)$ es un punto crítico de H para algunos valores $\lambda_i (i = 1, \ldots, k)$. Los parámetros λ_i se llaman multiplicadores de Lagrange.

Demostraremos que este método es válido para el caso especial en que $n = 3$ y $k = 1$.

7.3 Teorema. *Supongamos que F y G son funciones de* R^3 *en* R *que pertenecen a la clase C^1 en un conjunto abierto $\mathscr{E}$ y* $\mathbf{D}G$ *es distinta de cero en $\mathscr{E}$. Si $\mathbf{x}_0 = (x_0, y_0, z_0)$ es un punto en $\mathscr{E}$ en el que F tiene un valor extremo sujeto a la restricción $G(\mathbf{x}) = 0$, entonces, para algún valor de λ, (x_0, y_0, z_0, λ) es un punto crítico de*

$$H(x, y, z, \lambda) = F(x, y, z) + \lambda G(x, y, z).$$

PRUEBA. Supongamos que F restringido a la superficie $\mathscr{S}$ descrita por $G(\mathbf{x}) = 0$ tiene un valor extremo en el punto $\mathbf{x}_0$. Como $\mathbf{D}G(\mathbf{x}_0) \neq \mathbf{0}$, una de las derivadas parciales de G es distinta de cero en $\mathbf{x}_0$, digamos $D_3 G(\mathbf{x}_0) \neq 0$. De acuerdo con el teorema de la función implícita, pág. 229, existe una vecindad $\mathscr{N}$ de (x_0, y_0) y una función $g \in C^1$ sobre $\mathscr{N}$ tal que $g(x_0, y_0) = z_0$ y, para todo $(x, y) \in \mathscr{N}$,

$$G(x, y, g(x, y)) = 0$$

y

$$D_1 g(x, y) = -\frac{D_1 G(x, y, g(x, y))}{D_3 G(x, y, g(x, y))}$$

$$D_2 g(x, y) = -\frac{D_2 G(x, y, g(x, y))}{D_3 G(x, y, g(x, y))}.$$

Si $f(x, y) = F(x, y, g(x, y))$, entonces f tiene un valor extremo en (x_0, y_0) y

$$D_1 f(x_0, y_0) = D_1 F(x_0, y_0, z_0) + D_1 g(x_0, y_0) D_3 F(x_0, y_0, z_0)$$

$$= D_1 F(x_0, y_0, z_0) - \frac{D_1 G(x_0, y_0, z_0)}{D_3 G(x_0, y_0, z_0)} D_3 F(x_0, y_0, z_0) = 0$$

$$D_2 f(x_0, y_0) = D_2 F(x_0, y_0, z_0) + D_2 g(x_0, y_0) D_3 F(x_0, y_0, z_0)$$

$$= D_2 F(x_0, y_0, z_0) - \frac{D_2 G(x_0, y_0, z_0)}{D_3 G(x_0, y_0, z_0)} D_3 F(x_0, y_0, z_0) = 0 .$$

Luego, si $\lambda = -\dfrac{D_3 F(x_0, y_0, z_0)}{D_3 G(x_0, y_0, z_0)}$, tenemos

$$\begin{aligned}
D_1 H(x_0, y_0, z_0, \lambda) &= D_1 F(x_0, y_0, z_0) + \lambda D_1 G(x_0, y_0, z_0) = 0 \\
D_2 H(x_0, y_0, z_0, \lambda) &= D_2 F(x_0, y_0, z_0) + \lambda D_2 G(x_0, y_0, z_0) = 0 \\
D_3 H(x_0, y_0, z_0, \lambda) &= D_3 F(x_0, y_0, z_0) + \lambda D_3 G(x_0, y_0, z_0) = 0 \\
D_4 H(x_0, y_0, z_0, \lambda) &= G(x_0, y_0, z_0) = 0 .
\end{aligned}$$

Así pues, (x_0, y_0, z_0, λ) es un punto crítico de H. Lo que completa la prueba.

7.4 Ejemplo. (Véase pág. 243.) Una caja rectangular sin tapa ha de tener una superficie de área S. Encuéntrense las dimensiones de la caja que le darán el volumen máximo.

SOLUCIÓN. Supongamos que la caja tiene dimensiones x, y y z donde z es la altura. Deseamos encontrar el valor máximo de xyz cuando esta función está sujeta a la restricción $2xz+2yz+xy-S = 0$. Sea

$$H(x, y, z, \lambda) = xyz + \lambda(2xz+2yz+xy-S).$$

Para encontrar los puntos críticos de H debemos resolver las ecuaciones

7.5
$$\begin{aligned}
D_1 H(x, y, z, \lambda) &= yz + 2\lambda z + \lambda y = 0 \\
D_2 H(x, y, z, \lambda) &= xz + 2\lambda z + \lambda x = 0 \\
D_3 H(x, y, z, \lambda) &= xy + 2\lambda x + 2\lambda y = 0 \\
D_4 H(x, y, z, \lambda) &= 2xz + 2yz + xy - S = 0 .
\end{aligned}$$

De la primera ecuación obtenemos

7.6
$$2\lambda z = -yz - \lambda y .$$

Sustituyendo esto en la segunda ecuación tenemos

$$\begin{aligned}
xz - yz - \lambda y + \lambda x &= 0 \\
(\lambda + z)(x - y) &= 0 .
\end{aligned}$$

Así pues, $\lambda = -z$ o $x = y$. Si $\lambda = -z$, entonces la ecuación 7.6 implica que $z = 0$. Así pues, $\lambda = -z$ no da lugar al valor extremo deseado. Tomando $y = x$, la tercera ecuación de 7.5 implica que $\lambda = -\frac{1}{4}x$. Entonces, de la primera ecuación de 7.5 obtenemos $z = \frac{1}{2}x$. Así pues, la última ecuación

de 7.5 toma la forma $3x^2 = S$ y, por tanto, $x = \sqrt{\frac{S}{3}}$, $y = \sqrt{\frac{S}{3}}$, $z = \frac{1}{2}\sqrt{\frac{S}{3}}$.
Estas son las dimensiones de la caja de volumen máximo.

7.7 Ejemplo. Determínese el valor mínimo de z para puntos sobre la curva que es la intersección de la superficie descrita por $z = \sqrt{3x^2+8y^2+4}$ y el plano de ecuación $x+3y-z = 0$ (figura 8).

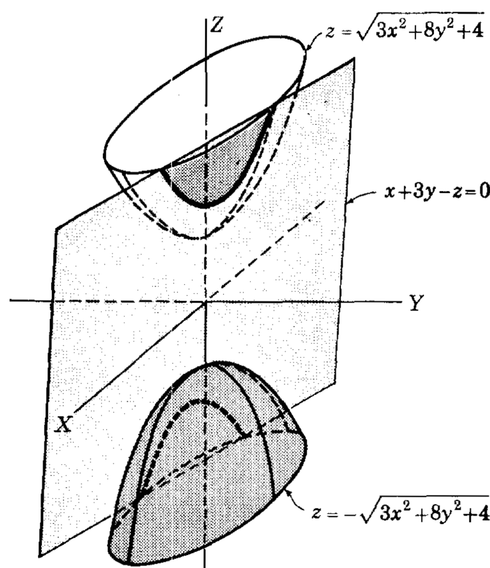


FIGURA 8

SOLUCIÓN. Deseamos encontrar el valor positivo mínimo de la función F definida por $F(x, y, z) = z$ sujeta a las restricciones $3x^2+8y^2-z^2+4 = 0$ y $x+3y-z = 0$. Sea

$$H(x, y, z, \lambda_1, \lambda_2) = z+\lambda_1(3x^2+8y^2-z^2+4)+\lambda_2(x+3y-z).$$

Para determinar los valores críticos de H debemos resolver las siguientes ecuaciones:

7.8
$$\begin{aligned} D_1 H(x, y, z, \lambda_1, \lambda_2) &= 6x\lambda_1+\lambda_2 = 0 \\ D_2 H(x, y, z, \lambda_1, \lambda_2) &= 16y\lambda_1+3\lambda_2 = 0 \\ D_3 H(x, y, z, \lambda_1, \lambda_2) &= 1-2z\lambda_1-\lambda_2 = 0 \\ D_4 H(x, y, z, \lambda_1, \lambda_2) &= 3x^2+8y^2-z^2+4 = 0 \\ D_5 H(x, y, z, \lambda_1, \lambda_2) &= x+3y-z = 0. \end{aligned}$$

Multiplicando la tres primeras ecuaciones por x, y y z, respectivamente, y sumando, obtenemos

$$2\lambda_1(3x^2+8y^2-z^2)+\lambda_2(x+3y-z)+z = 0.$$

Usando las últimas dos ecuaciones de 7.8, tenemos $z = 8\lambda_1$. Sustituyendo esto en la tercera ecuación de 7.8, obtenemos $\lambda_2 = 1 - 16\lambda_1^2$. Luego, según las dos primeras ecuaciones de 7.8, vemos que

$$x = \frac{16\lambda_1^2 - 1}{6\lambda_1} \quad \text{y} \quad y = \frac{48\lambda_1^2 - 3}{16\lambda_1}.$$

Sustituyendo estos valores de x, y y z en la última ecuación de 7.8, obtenemos $\lambda_1 = \pm \tfrac{1}{44}\sqrt{385}$. Como solamente estamos interesados en valores positivos de z, tomamos $z = \tfrac{2}{11}\sqrt{385}$. Basándose en consideraciones geométricas es fácil ver que éste es el valor mínimo deseado.

Problemas

1. Encuéntrense los valores extremos de las siguientes funciones sujetas a las restricciones que se indican:

a) $F(x, y, z) = x^2 + y^2 + z^2;\ \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} + 1 = 0$

b) $F(x, y, z) = x^2 + y^2 + z^2;\ \dfrac{x^2}{4} + \dfrac{y^2}{9} + z^2 = 1$

c) $F(x, y, z) = x + y + z;\ x^2 + y^2 + z^2 = 1$.

2. Úsense los multiplicadores de Lagrange para demostrar que la distancia más corta desde el punto (x_0, y_0, z_0) al plano $ax + by + cz + d = 0$ es la distancia perpendicular.

3. Determínese el paralelepípedo rectangular de volumen máximo con superficie de área igual a 48.

4. Determínese el paralelepípedo rectangular de volumen máximo que puede inscribirse en el elipsoide $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$.

5. Demuéstrese que si x, y, z son no negativos, entonces $F(x, y, z) = \sqrt[3]{xyz}$ con la restricción $x + y + z = 3a$ tiene un valor máximo cuando $x = y = z = a$. Demuéstrese que la media geométrica de tres números no negativos es igual o menor que su media aritmética, es decir, que

$$\sqrt[3]{xyz} \leqslant \tfrac{1}{3}(x + y + z).$$

6. Demuéstrese que si todos los x_i $(i = 1, \ldots, n)$ son no negativos, entonces

$$\left[\prod_{i=1}^{n} x_i\right]^{1/n} \leqslant \frac{1}{n}\sum_{i=1}^{n} x_i.$$

7. Encuéntrense los puntos más cercanos al origen de la intersección del hiperboloide de una hoja $x^2+y^2-z^2 = 1$ y el plano $2x+y+z = 0$.

8. INTEGRALES CURVILÍNEAS

En esta sección discutiremos un tipo de integral que se usa en muchas aplicaciones físicas. Es ésta una integral de una función vectorial de un vector a lo largo de cierta curva en el dominio de la función. Para simplificar la discusión de tales integrales, nos restringiremos a la consideración de funciones y curvas que son del tipo que con más frecuencia ocurren en las aplicaciones físicas.

Sea $\mathscr{C}$ la curva descrita por la transformación $\mathbf{x}$ de $[a, b]$ una curva lisa en R^n; es decir, supongamos que $\mathbf{x}'$ es continua y distinta de cero sobre $[a, b]$. Sea $\mathbf{f}$ una función de R^n en R^n que es continua sobre un conjunto abierto que contiene a $\mathscr{C}$.

8.1 Definición. *La **integral curvilínea** de* $\mathbf{f}$ *a lo largo de la curva lisa* $\mathscr{C}$ *descrita por el mapeo* $\mathbf{x}$ *de* $[a, b]$ *se denota por* $\int_{\mathscr{C}} \mathbf{f}\cdot d\mathbf{x}$ *y se define por*

$$\int_{\mathscr{C}} \mathbf{f}\cdot d\mathbf{x} = \int_a^b \mathbf{f}(\mathbf{x}(t))\cdot\mathbf{x}'(t)\,dt.$$

Como supusimos que $\mathbf{x}'$ es continua sobre $[a, b]$ y que $\mathbf{f}$ es continua sobre $\mathscr{C}$, la integral que aparece en el segundo miembro existe.

8.2 Ejemplo. Evalúese la integral curvilínea $\int_{\mathscr{C}} (x^2, 2xy)\cdot d\mathbf{x}$ donde $\mathscr{C}$ es la semicircunferencia descrita por $x = \cos t$, $y = \operatorname{sen} t$, $t\in[0, \pi]$.

SOLUCIÓN.

$$\int_{\mathscr{C}} (x, 2xy)\cdot d\mathbf{x} = \int_0^\pi (\cos^2 t, 2\cos t \operatorname{sen} t)\cdot(-\operatorname{sen} t, \cos t)\,dt$$

$$= \int_0^\pi \cos^2 t \operatorname{sen} t\,dt$$

$$= -\tfrac{1}{3}\cos^3 t\Big]_0^\pi = \tfrac{2}{3}.$$

8.3 Ejemplo. Evalúese la integral curvilínea $\int_{\mathscr{C}} (y, x)\cdot d\mathbf{x}$ donde $\mathscr{C}$ es el arco de la parábola $y = x^2$ desde $(0, 0)$ hasta $(2, 4)$.

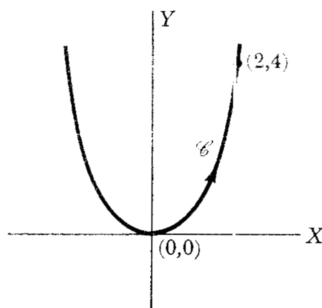


FIGURA 9

SOLUCIÓN. La curva $\mathscr{C}$ (figura 9) tiene una representación paramétrica $x = t$, $y = t^2$, $t \in [0, 2]$.
Así pues,

$$\int_{\mathscr{C}} (y, x) \cdot d\mathbf{x} = \int_0^2 (t^2, t) \cdot (1, 2t)\, dt$$

$$= \int_0^2 3t^2\, dt = t^3 \Big]_0^2 = 8.$$ [1]

La integral curvilínea de una función $\mathbf{f} = (f_1, f_2)$ a lo largo de una curva $\mathscr{C}$ se escribe frecuentemente $\int_{\mathscr{C}} f_1\, dx + f_2\, dy$. Si hacemos $d\mathbf{x} = (dx, dy)$, entonces es natural escribir

$$\int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} = \int_{\mathscr{C}} (f_1, f_2) \cdot (dx, dy) = \int_{\mathscr{C}} f_1\, dx + f_2\, dy.$$

Además, si la curva $\mathscr{C}$ está descrita por las ecuaciones $x = x(t)$, $y = y(t)$, $t \in [a, b]$, entonces la integral curvilínea de f a lo largo de $\mathscr{C}$ está definida por la integral de Riemann

$$\int_a^b \mathbf{f}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\, dt = \int_a^b [f_1(x(t), y(t))\, x'(t) + f_2(x(t), y(t))\, y'(t)]\, dt.$$

Así pues, en $\int_{\mathscr{C}} f_1\, dx + f_2\, dy$ podemos considerar dx y dy como diferenciales, y esta notación nos guía en una evaluación correcta de la integral curvilínea.

[1] Obsérvese que no se ha probado la independencia de la representación paramétrica. [N. del T.]

Así, en el ejemplo 8.3 deseamos evaluar la integral curvilínea $\int_{\mathscr{C}} y\,dx + x\,dy$ donde $\mathscr{C}$ está descrita por $x = t$, $y = t^2$, $t \in [0, 2]$.

$$\int_{\mathscr{C}} y\,dx + x\,dy = \int_0^2 (t^2 + 2t^2)\,dt = \int_0^2 3t^2\,dt = 8 .$$

De la definición de integral curvilínea y de las propiedades de la integral de Riemann es fácil deducir que

$$\int_{\mathscr{C}} c\mathbf{f} \cdot d\mathbf{x} = c \int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x}$$

$$\int_{\mathscr{C}} (\mathbf{f} + \mathbf{g}) \cdot d\mathbf{x} = \int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} + \int_{\mathscr{C}} \mathbf{g} \cdot d\mathbf{x} .$$

Si $\mathscr{C}$ es la curva lisa descrita por la ecuación $\mathbf{x} = \mathbf{g}(t)$, $t \in [a, b]$, entonces denotamos por $-\mathscr{C}$ la curva "recorrida" en dirección opuesta a la de la $\mathscr{C}$; es decir, $-\mathscr{C}$ está descrita por $\mathbf{x} = \mathbf{g}(-t)$, $t \in [-b, -a]$. Entonces,

$$\int_{-\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} = - \int_{-b}^{-a} \mathbf{f}(\mathbf{g}(-t)) \cdot \mathbf{g}'(-t)\,dt .$$

Haciendo $u = -t$ obtenemos

$$\int_{-\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} = \int_b^a \mathbf{f}(\mathbf{g}(u)) \cdot \mathbf{g}'(u)\,du = - \int_a^b \mathbf{f}(\mathbf{g}(u)) \cdot \mathbf{g}'(u)\,du$$

$$= - \int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x}.^{1}$$

Además, si la curva $\mathscr{C}$ está compuesta de las curvas $\mathscr{C}_1$ y $\mathscr{C}_2$ (es decir, si $\mathscr{C}$ se traza primero $\mathscr{C}_1$ y después $\mathscr{C}_2$), entonces

$$\int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} = \int_{\mathscr{C}_1} \mathbf{f} \cdot d\mathbf{x} + \int_{\mathscr{C}_2} \mathbf{f} \cdot d\mathbf{x} .$$

[1] Quizá convenga precisar un poco más. Si $\mathscr{C}$ es la curva dada por $\mathbf{g}$: $[a, b] \to \mathrm{R}^n$, entonces $-\mathscr{C}$ es la curva dada por $\mathbf{h}$: $[-b, -a] \to \mathrm{R}^n$ tal que $\mathbf{h}(r) = \mathbf{g}(-r)$. De acuerdo con las definiciones dadas y el cambio de variable r: $[-b, -a] \to [a, b]$ tal que $r(t) = -t$, se tiene:

$$\int_{-\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} = \int_{-b}^{-a} \mathbf{f}(\mathbf{h}(r)) \cdot \mathbf{h}'(r)\,dr = - \int_{-b}^{-a} \mathbf{f}(\mathbf{g}(-r)) \cdot \mathbf{g}'(-r)\,dr$$

$$= - \int_b^a \mathbf{f}(\mathbf{g}(t)) \cdot \mathbf{g}'(t)\,(-dt) = \int_b^a \mathbf{f}(\mathbf{g}(t)) \cdot \mathbf{g}'(t)\,dt$$

$$= - \int_a^b \mathbf{f}(\mathbf{g}(t)) \cdot \mathbf{g}'(t)\,dt = - \int_{\mathscr{C}}^{b} \mathbf{f} \cdot d\mathbf{x}. \text{ [N. del T.]}$$

Supongamos que $\mathscr{C}$ está descrito por la transformación $\mathbf{x}$ de $[a, b]$ y $\mathscr{C}_1$ y $\mathscr{C}_2$ están descritos por la transformación $\mathbf{x}$ de $[a, c]$ y $[c, b]$, respectivamente, donde $c \in \langle a, b \rangle$. Entonces

$$\begin{aligned}\int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} &= \int_a^b \mathbf{f}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\, dt \\ &= \int_a^c \mathbf{f}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\, dt + \int_c^b \mathbf{f}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\, dt \\ &= \int_{\mathscr{C}_1} \mathbf{f} \cdot d\mathbf{x} + \int_{\mathscr{C}_2} \mathbf{f} \cdot d\mathbf{x}.\end{aligned}$$

Extendemos ahora la definición de integral curvilínea a una trayectoria compuesta de cierto número de curvas lisas que no necesariamente han de formar una curva lisa. Una trayectoria consistente en un número finito de curvas lisas se llama *curva lisa en trozos*. Si la curva lisa en trozos $\mathscr{C}$ se compone de las curvas lisas $\mathscr{C}_1, \ldots, \mathscr{C}_m$, entonces definimos

$$\int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} = \sum_{k=1}^{m} \int_{\mathscr{C}_k} \mathbf{f} \cdot d\mathbf{x}.$$

8.4 Ejemplo. Evalúese la integral curvilínea $\int_{\mathscr{C}} 2xy\, dx + x^2\, dy$ donde $\mathscr{C}$ es la frontera del triángulo con vértices $(0, 0)$, $(3, 2)$ y $(1, 5)$ recorrida en dirección levógira (figura 10).

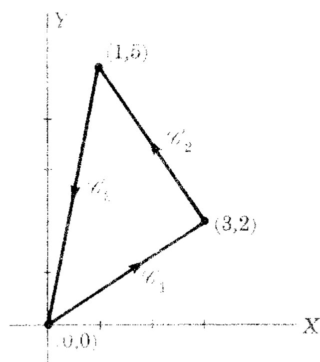


FIGURA 10

SOLUCIÓN. Sean $\mathscr{C}_1$, $\mathscr{C}_2$ y $\mathscr{C}_3$ los segmentos rectilíneos de $(0, 0)$ a $(3, 2)$, de $(3, 2)$ a $(1, 5)$, y de $(1, 5)$ a $(0, 0)$, respectivamente. $\mathscr{C}$ es lisa en trozos e

$$\int_{\mathscr{C}} 2xy\, dx + x^2\, dy = \sum_{k=1}^{3} \int_{\mathscr{C}_k} 2xy\, dx + x^2\, dy.$$

$\mathscr{C}_1$ tiene como ecuación $y = \frac{2}{3}x$ donde x va de 0 a 3. Podemos usar x como un parámetro para esta curva: $x = x$, $y = \frac{2}{3}x$, $x \in [0, 3]$. Entonces

$$\int_{\mathscr{C}_1} 2xy\,dx + x^2\,dy = \int_0^3 (\tfrac{4}{3}x^2 + \tfrac{2}{3}x^2)\,dx = 18\,.$$

La curva $-\mathscr{C}_2$ tiene la representación paramétrica: $x = x$, $y = 5 - \frac{3}{2}(x-1)$, $x \in [1, 3]$. Entonces

$$\int_{\mathscr{C}_2} 2xy\,dx + x^2\,dy = -\int_{-\mathscr{C}_2} 2xy\,dx + x^2\,dy$$

$$= -\int_1^3 \{2x[5 - \tfrac{3}{2}(x-1)] + x^2(-\tfrac{3}{2})\}\,dx$$

$$= \int_3^1 (13x - \tfrac{9}{2}x^2)\,dx = -13\,.$$

$\mathscr{C}_3$ tiene como ecuación $y = 5x$ donde x va de 1 a 0. Por tanto

$$\int_{\mathscr{C}_3} 2xy\,dx + x^2\,dy = \int_1^0 (10x^2 + 5x^2)\,dx = -5\,.$$

Luego

$$\int_{\mathscr{C}} 2xy\,dx + x^2\,dy = 0\,.$$

Probaremos ahora un teorema para integrales curvilíneas que es análogo al segundo teorema fundamental del cálculo.

8.5 Teorema. *Supongamos que* $\mathbf{f}$ *es continua sobre un conjunto abierto* $\mathscr{E}$ *y sean* $\mathbf{x}_1$ *y* $\mathbf{x}_2$ *dos puntos cualesquiera de* $\mathscr{E}$. *Entonces, si* $\mathbf{f} = \mathbf{D}g$ *sobre* $\mathscr{E}$ *y si* $\mathscr{C}$ *es una curva lisa en trozos en* $\mathscr{E}$ *de* $\mathbf{x}_1$ *a* $\mathbf{x}_2$,

$$\int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} = g(\mathbf{x}_2) - g(\mathbf{x}_1)\,.$$

PRUEBA. Supongamos que $\mathscr{C}$ está descrita por la transformación $\mathbf{x}$ de $[a, b]$ y sea $h(t) = g(\mathbf{x}(t))$. Entonces $h'(t) = \mathbf{D}g(\mathbf{x}(t)) \cdot \mathbf{x}'(t)$ e

$$\int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} = \int_{\mathscr{C}} \mathbf{D}g \cdot d\mathbf{x} = \int_a^b \mathbf{D}g(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\,dt$$

$$= \int_a^b h'(t)\,dt = h(b) - h(a) = g(\mathbf{x}_2) - g(\mathbf{x}_1)\,.$$

Nota. En la prueba de este teorema hemos aplicado el segundo teorema fundamental del cálculo a la función h' que es continua a trozos sobre $[a, b]$. Decimos que una función es continua a trozos sobre $[a, b]$ si la función es continua en todos, salvo en un número finito de puntos de $[a, b]$ y en cada punto de discontinuidad existen los límites a la derecha y a la izquierda de la función. Aunque el segundo teorema fundamental se enuncia generalmente para funciones con derivada continua, la extensión al caso en que la derivada es continua a trozos no presenta dificultad alguna.

El teorema 8.5 enuncia que la integral curvilínea de una función que es la derivada de alguna función g depende solamente de los valores de g en los extremos $\mathbf{x}_1$ y $\mathbf{x}_2$ de la curva. Así pues, la integral curvilínea de una función tal es independiente de la curva lisa a trozos en $\mathscr{E}$ que una $\mathbf{x}_1$ y $\mathbf{x}_2$: decimos por ello que la integral curvilínea es *independiente de la trayectoria* en $\mathscr{E}$.

Nota. Definimos una curva cerrada como una cuyos extremos coinciden. Es claro que una integral curvilínea es independiente de la trayectoria en $\mathscr{E}$ si y sólo si la integral curvilínea a lo largo de cualquier curva cerrada lisa a trozos en $\mathscr{E}$ es cero.

8.6 Ejemplo. Evalúese $\int_{\mathscr{C}} y\,dx+(x+1)\,dy$ donde $\mathscr{C}$ es el arco de la elipse $9x^2+4y^2 = 36$ desde $(2, 0)$ hasta $(0, 3)$.

SOLUCIÓN. Usaremos el teorema 8.5 para evaluar esta integral. Si $(y, x+1)$ es la derivada de una función g, entonces debemos tener $D_1 g(x, y) = y$ y $D_2 g(x, y) = x+1$. Considerando $D_1 g(x, y) = y$ como una ecuación diferencial en la que y es una constante, vemos que

8.7 $$g(x, y) = xy+\varphi(y).$$

Además, $D_2 g(x, y) = x+1$ implica que

8.8 $$g(x, y) = xy+y+\psi(x).$$

Así pues, si tomamos $\varphi(y) = y$ y $\psi(x) = 0$ y, por tanto, $g(x, y) = xy+y$, vemos que tanto 8.7 como 8.8 se satisfacen. Es fácil verificar que para esta función g

$$\mathbf{D}g(x, y) = (y, x+1).$$

Luego

$$\int_{\mathscr{C}} y\,dx+(x+1)\,dy = g(0,3)-g(2,0) = 3.$$

Tenemos el siguiente corolario al teorema 8.5.

8.9 Corolario. *Supongamos que* $\mathbf{f}$ *es continua sobre un conjunto abierto* $\mathscr{E}$ *y* $\mathbf{f} = \mathbf{D}g$ *sobre* $\mathscr{E}$. *Entonces, si* $\mathscr{C}$ *es una curva cerrada lisa a trozos en* $\mathscr{E}$,

$$\int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} = 0 .$$

PRUEBA. En el teorema 8.5 no exigimos que los puntos extremos $\mathbf{x}_1$ y $\mathbf{x}_2$ fueran distintos. Si $\mathscr{C}$ es una curva cerrada, entonces los puntos extremos coinciden y tenemos

$$\int_{\mathscr{C}} \mathbf{f} \cdot d\mathbf{x} = g(\mathbf{x}_1) - g(\mathbf{x}_1) = 0 .$$

Podemos usar este corolario para demostrar que $\int_{\mathscr{C}} 2xy\,dx + x^2\,dy = 0$ donde $\mathscr{C}$ es la frontera del triángulo con vértices $(0, 0)$, $(3, 2)$ y $(1, 5)$ (ejemplo 8.4). Si $(2xy, x^2)$ es la derivada de una función g, entonces $D_1 g(x, y) = 2xy$ y $D_2 g(x, y) = x^2$. La ecuación $D_1 g(x, y) = 2xy$ implica que

$$g(x, y) = x^2 y + \varphi(y)$$

y $D_2 g(x, y) = x^2$ implica que

$$g(x, y) = x^2 y + \psi(x) .$$

Si tomamos $\varphi(y) = 0 = \psi(x)$, entonces $g(x, y) = x^2 y$. Se verifica fácilmente que la derivada de g es $(2xy, x^2)$. Por tanto

$$\int_{\mathscr{C}} 2xy\,dx + x^2\,dy = 0 .$$

No toda función continua de R^n en R^n es la derivada de una función de R^n en R. Consideremos, por ejemplo, la función $\mathbf{f}$ definida por $\mathbf{f}(x, y) = (2xy, x^3)$. Si $\mathbf{f}$ fuera la derivada de una función g, entonces $D_1 g(x, y) = 2xy$ y $D_2 g(x, y) = x^3$. La ecuación $D_1 g(x, y) = 2xy$ implica que

8.10 $$g(x, y) = x^2 y + \varphi(y)$$

y $D_2 g(x, y) = x^3$ implica que

8.11 $$g(x, y) = x^3 y + \psi(x) .$$

Claramente no hay función alguna que pueda satisfacer simultáneamente a 8.10 y 8.11, por tanto, $\mathbf{f}$ no es la derivada de función alguna.

La expresión $\mathbf{f} \cdot d\mathbf{x}$ se llama *diferencial exacta* sobre un conjunto abierto $\mathscr{E}$ de R^n si hay una función g de R^n en R tal que $\mathbf{f} = \mathbf{D}g$ sobre $\mathscr{E}$, y, por tanto,

$$\mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} = \mathbf{D}g(\mathbf{x}) \cdot d\mathbf{x} = dg(\mathbf{x}; d\mathbf{x}) .$$

El teorema 8.5 muestra que si $\mathbf{f}\cdot d\mathbf{x}$ es una diferencial exacta sobre $\mathscr{E}$, entonces la integral curvilínea $\int_{\mathscr{C}} \mathbf{f}\cdot d\mathbf{x}$ es independiente de la trayectoria en $\mathscr{E}$.

Si $\mathbf{f}\in C^1$ sobre $\mathscr{E}$ y existe una función g tal que $\mathbf{f} = \mathbf{D}g$ sobre $\mathscr{E}$, entonces $g\in C^2$ sobre $\mathscr{E}$ y, por tanto, de acuerdo con el teorema 10.3, pág. 214,

$$D_{i,j}g = D_{j,i}g \quad \text{sobre } \mathscr{E}\,(i, j = 1, \ldots, n)$$

o, lo que es equivalente,

8.12 $$D_i f_j = D_j f_i \quad \text{sobre } \mathscr{E}\,(i, j = 1, \ldots, n).$$

Esto nos da una condición necesaria para que $\mathbf{f}\cdot d\mathbf{x}$ sea una diferencial exacta sobre $\mathscr{E}$. Así pues, si $D_i f_j(\mathbf{x}) \neq D_j f_i(\mathbf{x})$ para algún $\mathbf{x}\in\mathscr{E}$ y algún i y j, entonces $\mathbf{f}$ no es la derivada de una función sobre $\mathscr{E}$. Por ejemplo, si $\mathbf{f}(x, y) = (2xy, x^3)$, entonces

$$D_1 f_2(x, y) = 3x^2 \quad \text{y} \quad D_2 f_1(x, y) = 2x.$$

Como estas derivadas parciales no son iguales no hay función alguna g tal que $\mathbf{D}g = \mathbf{f}$, es decir, $\mathbf{f}\cdot d\mathbf{x}$ no es una diferencial exacta.

La continuidad y la igualdad de las derivadas parciales (8.12) no son suficientes para poder asegurar que $\mathbf{f}\cdot d\mathbf{x}$ es una diferencial exacta sobre $\mathscr{E}$. Debemos poner algunas restricciones sobre el conjunto abierto $\mathscr{E}$ (véase el teorema 13.8, pág. 679).

Además, el recíproco del teorema 8.5 no se verifica a menos que se establezcan algunas restricciones sobre $\mathscr{E}$. Decimos que un conjunto $\mathscr{E}$ es *arcoconectable* si para cualesquier dos puntos $\mathbf{x}_1$ y $\mathbf{x}_2$ de $\mathscr{E}$ hay una curva lisa a trozos en $\mathscr{E}$ con $\mathbf{x}_1$ y $\mathbf{x}_2$ como puntos extremos. Se puede demostrar que si un conjunto es arcoconectable, entonces es conexo (problema 10, pág. 254). El recíproco no es en general cierto. Sin embargo, si $\mathscr{E}$ es abierto y conexo, entonces $\mathscr{E}$ es arcoconectable. Para una discusión del concepto de conexidad véase la referencia (2) o la (7).

8.13 Teorema. *Si* $\mathbf{f}$ *es continua sobre un conjunto abierto conexo* $\mathscr{E}$ *y la integral de curva de* $\mathbf{f}$ *es independiente de la trayectoria en* $\mathscr{E}$, *entonces* $\mathbf{f}\cdot d\mathbf{x}$ *es una diferencial exacta en* $\mathscr{E}$.

PRUEBA. Escojamos algún punto $\mathbf{x}_0$ en $\mathscr{E}$. Para cualquier punto $\mathbf{x}\in\mathscr{E}$, sea

$$g(\mathbf{x}) = \int_{\mathscr{C}} \mathbf{f}\cdot d\mathbf{x}$$

donde $\mathscr{C}$ es una curva lisa a trozos que va de $\mathbf{x}_0$ a $\mathbf{x}_1$ y toda ella situada en $\mathscr{E}$. Como $\mathscr{E}$ es abierto y conexo una curva tal existe. Además, como la integral curvilínea de $\mathbf{f}$ es independiente de la trayectoria en $\mathscr{E}$, el valor

$g(\mathbf{x})$ no depende de la elección de la curva $\mathscr{C}$. Consideremos ahora un punto particular $\mathbf{x}$ en $\mathscr{E}$ y sea $\mathscr{C}_1$ una curva lisa a trozos de $\mathbf{x}_0$ a $\mathbf{x}$ que pasa por $\mathscr{E}$. Como $\mathscr{E}$ es abierto, hay una vecindad $\mathscr{N}(\mathbf{x};\delta)$ de $\mathbf{x}$ contenida en $\mathscr{E}$. Entonces, para $|h| < \delta$ el segmento rectilíneo

$$\mathscr{C}_2 = \{\mathbf{x}+th\mathbf{u}_k \mid t\in[0,1]\},$$

donde $\mathbf{u}_k$ denota el vector unitario en la dirección del eje X_k, se encuentra en $\mathscr{E}$. Sea $\mathscr{C}_3$ la trayectoria compuesta de $\mathscr{C}_1$ y $\mathscr{C}_2$. Tenemos entonces

$$g(\mathbf{x}+h\mathbf{u}_k)-g(\mathbf{x}) = \int_{\mathscr{C}_3} \mathbf{f}\cdot d\mathbf{x} - \int_{\mathscr{C}_1} \mathbf{f}\cdot d\mathbf{x} = \int_{\mathscr{C}_2} \mathbf{f}\cdot d\mathbf{x}$$

$$= \int_0^1 \mathbf{f}(\mathbf{x}+th\mathbf{u}_k)\cdot h\mathbf{u}_k\, dt = h\int_0^1 f_k(\mathbf{x}+th\mathbf{u}_k)\, dt$$

$$= hf_k(\mathbf{x}+\theta h\mathbf{u}_k) \quad \text{para algún } \theta\in\langle 0,1\rangle,$$

donde el último paso se obtuvo al aplicar el primer teorema del valor medio para integrales. Así pues, como $\mathbf{f}$ es continua sobre $\mathscr{E}$,

$$\lim_{h\to 0} \frac{g(\mathbf{x}+h\mathbf{u}_k)-g(\mathbf{x})}{h} = \lim_{h\to 0} f_k(\mathbf{x}+\theta h\mathbf{u}_k) = f_k(\mathbf{x});$$

es decir, $D_k g(\mathbf{x}) = f_k(\mathbf{x})$. Esto demuestra que $\mathbf{D}g = \mathbf{f}$ sobre $\mathscr{E}$ y, por tanto, $\mathbf{f}\cdot d\mathbf{x}$ es una diferencial exacta sobre $\mathscr{E}$.

Problemas

1. Evalúense las siguientes integrales curvilíneas:

a) $\int_{\mathscr{C}} (xy^2, x)\cdot d\mathbf{x}$ donde $\mathscr{C}$ es el arco de la elipse:

$x = \cos t,\ y = 3 \operatorname{sen} t,\ t\in[0,\pi]$

b) $\int_{\mathscr{C}} (y, x)\cdot d\mathbf{x}$ donde $\mathscr{C}$ es el arco de la hipérbola:

$x = \cosh t,\ y = \operatorname{senh} t,\ t\in[-1,2]$

c) $\int_{\mathscr{C}} (x, y)\cdot d\mathbf{x}$ donde $\mathscr{C}$ es el arco de la parábola:

$x = t^2,\ y = t,\ t\in[-2,2]$

d) $\int_{\mathscr{C}} (xy, z^2, y)\cdot d\mathbf{x}$ donde $\mathscr{C}$ es el arco de la hélice cilíndrica:

$x = \cos t,\ y = \operatorname{sen} t,\ z = t,\ t\in[0,2\pi]$

e) $\int_{\mathscr{C}} (x^2+y, yz, x+3y)\cdot d\mathbf{x}$ donde $\mathscr{C}$ es el arco de la cúbica alabeada: $x = t,\ y = t^2,\ z = t^3,\ t\in[0,2]$

f) $\int_{\mathscr{C}} 3xy^2\,dx+(x+3y)\,dy$ donde $\mathscr{C}$ es el segmento rectilíneo: $x = 3t+2,\ \ y = t-1,\ t\in[-1,3]$

g) $\int_{\mathscr{C}} (x^3-xy^2)\,dx$ donde $\mathscr{C}$ es la curva: $x = t^2,\ \ y = t^3,\ t\in[0,2]$

h) $\int_{\mathscr{C}} xy\,dx+z\,dy+xz\,dz$ donde $\mathscr{C}$ es la curva: $x = t,\ \ y = t,\ \ z = t^2,\ t\in[-3,3]$.

2. Evalúense las siguientes integrales curvilíneas:

a) $\int_{\mathscr{C}} y^2\,dx+x^2\,dy$ donde $\mathscr{C}$ es el segmento de la recta $y = 2x+3$ de $(-1,1)$ a $(2,7)$

b) $\int_{\mathscr{C}} x\,dx+y\,dy$ donde $\mathscr{C}$ es el arco de la curva $y = x^3$ de $(2,8)$ a $(0,0)$

c) $\int_{\mathscr{C}} (x+y)\,dx+x^2\,dy$ donde $\mathscr{C}$ es el arco de la curva $y = \operatorname{sen} x$ de $(0,0)$ a $(\pi/2,1)$

d) $\int_{\mathscr{C}} 3xy\,dx+(xy^2+y)\,dy$ donde $\mathscr{C}$ es el arco de la curva $x-y^4 = 0$ desde $(1,1)$ a $(0,0)$.

3. Evalúense las integrales curvilíneas $\int_{\mathscr{C}} (y+3)\,dx+(x-2)\,dy$ e

$$\int_{\mathscr{C}} xy\,dx+dy \text{ cuando}$$

a) $\mathscr{C}$ es la frontera del rectángulo con vértices (0, 0), (2, 0), (2, 1) y (0, 1) recorrida en dirección levógira.

b) $\mathscr{C}$ es la circunferencia $x = \cos t,\ y = \operatorname{sen} t,\ t\in[0, 2\pi]$.

4. Determínese si sí o no $\mathbf{f}\cdot d\mathbf{x}$ es una diferencial exacta:

a) $\mathbf{f}(x, y) = (x+3, y-2)$
b) $\mathbf{f}(x, y) = (y-2, 2x)$
c) $\mathbf{f}(x, y, z) = (x+y, 2z, yz)$
d) $\mathbf{f}(x, y, z) = (2xz+2y, 2x, x^2+3)$.

5. Evalúese la integral $\int_{\mathscr{C}} (2xy^3+2x)dx+(3x^2y^2+1)dy$, cuando $\mathscr{C}$ es el arco de la cicloide: $x = t - \operatorname{sen} t$, $y = 1 - \cos t$, $t \in [0, 2\pi]$.

6. Evalúese $\int_{\mathscr{C}} \frac{-y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy$ cuando

a) $\mathscr{C}$ es la circunferencia unitaria: $x = \cos t$, $y = \operatorname{sen} t$, $t \in [0, 2\pi]$,

b) $\mathscr{C}$ es la circunferencia unitaria recorrida dos veces: $x = \cos t$, $y = \operatorname{sen} t$, $t \in [0, 4\pi]$,

c) $\mathscr{C}$ es la circunferencia: $x = 3 + \cos t$, $y = 2 + \operatorname{sen} t$, $t \in [0, 2\pi]$.

9. APLICACIONES A LA MECÁNICA

Supongamos que **F** es un campo de fuerzas tridimensional; es decir, **F** es una función que asigna a cada punto **x** de alguna región $\mathscr{E}$ de R^3 la fuerza **F**(**x**) que actuaría sobre una partícula en este punto. Deseamos definir el trabajo hecho por el campo de fuerzas al mover una partícula a lo largo de una curva $\mathscr{C}$ en $\mathscr{E}$. El concepto básico de trabajo hecho por una fuerza al mover una partícula de una posición a otra es el componente de la fuerza en la dirección del movimiento por la distancia recorrida. Sea $\mathscr{C}$ una curva lisa descrita por la ecuación $\mathbf{x} = \mathbf{x}(t)$, $t \in [a, b]$. En el punto $\mathbf{x}(t)$ el componente de la fuerza en la dirección del movimiento es $\mathbf{F}(\mathbf{x}(t)) \cdot \frac{\mathbf{x}'(t)}{|\mathbf{x}'(t)|}$ donde $\frac{\mathbf{x}'(t)}{|\mathbf{x}'(t)|}$ es un vector tangente unitario en la dirección del crecimiento del parámetro. Así pues, si tomamos una partición $\{t_i \mid i = 0, \ldots, n\}$ del intervalo $[a, b]$, el trabajo hecho por el campo de fuerza al mover una partícula a lo largo de $\mathscr{C}$ sería aproximadamente $\sum_{i=1}^{n} \mathbf{F}(\mathbf{x}(\bar{t}_i)) \cdot \mathbf{x}'(\bar{t}_i) \Delta_i t$ donde $\Delta_i t = t_i - t_{i-1}$ y $\bar{t}_i \in [t_{i-1}, t_i]$. Si estas sumas aproximativas tienden a un número cuando la norma de la partición tiende a cero, entonces ese límite es, por definición, el trabajo hecho por el campo de fuerzas. Suponiendo que $\mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)$ es continua a trozos, sabemos que tal límite existe y es la integral

$$\int_a^b \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\,dt = \int_{\mathscr{C}} \mathbf{F} \cdot d\mathbf{x}.$$

Así pues, *el trabajo hecho por un campo de fuerzas* **F** *al mover una partícula a lo largo de una curva* $\mathscr{C}$ *es, por definición, la integral curvilínea* $\int_{\mathscr{C}} \mathbf{F} \cdot d\mathbf{x}$.

Nota. Para asegurarnos de la existencia de la integral curvilínea $\int_{\mathscr{C}} \mathbf{F} \cdot d\mathbf{x}$

suqondremos a lo largo de toda esta sección que $\mathbf{F}$ siempre denota una función continua definida sobre un conjunto abierto $\mathscr{E}$ y $\mathscr{C}$ es una curva lisa a trozos contenida en $\mathscr{E}$.

Si $\mathbf{x}_1$ y $\mathbf{x}_2$ son los puntos extremos de la curva $\mathscr{C}$, entonces denotaremos la integral curvilínea de $\mathbf{F}$ a lo largo de $\mathscr{C}$ por $\int_{\mathscr{C}\,\mathbf{x}_1}^{\mathbf{x}_2} \mathbf{F}\cdot d\mathbf{x}$.

La segunda ley del movimiento de Newton afirma que una partícula de masa m sujeta a un campo de fuerza $\mathbf{F}$ se moverá de acuerdo a la ecuación

$$m\mathbf{x}(t) = \mathbf{F}(\mathbf{x}(t))$$

donde $\mathbf{x}(t)$ es la posición de la partícula en el instante t. Entonces

$$m\mathbf{x}(t)\cdot\dot{\mathbf{x}}(t) = \tfrac{1}{2}mD_t[\dot{\mathbf{x}}(t)\cdot\dot{\mathbf{x}}(t)] = \mathbf{F}(\mathbf{x}(t))\cdot\dot{\mathbf{x}}(t)$$

y

9.1 $$\tfrac{1}{2}m|\mathbf{v}(t_2)|^2 - \tfrac{1}{2}m|\mathbf{v}(t_1)|^2 = \int_{t_1}^{t_2} \mathbf{F}(\mathbf{x}(t))\cdot\dot{\mathbf{x}}(t)\,dt = \int_{\mathscr{C}\,\mathbf{x}(t_1)}^{\mathbf{x}(t_2)} \mathbf{F}\cdot d\mathbf{x}.$$

La cantidad $\frac{1}{2}m|\mathbf{v}(t)|^2$ se llama *energía cinética* de la partícula en el instante t. Por tanto, 9.1 nos dice: *cuando una partícula se mueve a lo largo de su trayectoria $\mathscr{C}$ desde $\mathbf{x}(t_1)$ hasta $\mathbf{x}(t_2)$, el cambio en energía cinética es igual al trabajo hecho por el campo de fuerzas.*

Un caso particularmente importante se presenta cuando el campo de fuerzas es conservativo. Un campo de fuerzas $\mathbf{F}$ definido sobre $\mathscr{E}$ se dice que es *conservativo* si el trabajo realizado a lo largo de toda curva cerrada es cero. Esto es equivalente a decir que el campo de fuerzas $\mathbf{F}$ es conservativo si el trabajo hecho al mover una partícula de una posición a otra es independiente de la trayectoria a lo largo de la cual la partícula se haya movido. Si $\mathbf{F}$ es conservativo y el dominio de $\mathbf{F}$ es un conjunto abierto y conexo $\mathscr{E}$, entonces el teorema 8.13 implica que existe una función real U definida sobre $\mathscr{E}$, llamada *función potencial*, tal que $\mathbf{D}U = -\mathbf{F}$ sobre $\mathscr{E}$. Además, según el teorema 8.5, vemos que si $\mathbf{F}$ tiene una función potencial U entonces $\mathbf{F}$ es conservativo e

$$\int_{\mathscr{C}\,\mathbf{x}_1}^{\mathbf{x}_2} \mathbf{F}\cdot d\mathbf{x} = U(\mathbf{x}_1) - U(\mathbf{x}_2).$$

Así pues, cuando el campo de fuerza es conservativo, la ecuación 9.1 puede escribirse como

$$\tfrac{1}{2}m|v(t_2)|^2 + U(\mathbf{x}(t_2)) = \tfrac{1}{2}m|\mathbf{v}(t_1)|^2 + U(\mathbf{x}(t_1)).$$

Esta es la ley de la conservación de la energía: *si el campo de fuerza es conservativo, la suma de la energía cinética y de la energía potencial es una constante.*

Si U es una función potencial para $\mathbf{F}$, entonces $\mathbf{F} = -\mathbf{D}U$. Esto implica que en un punto $\mathbf{x}$ la fuerza es ortogonal a la superficie que pasa por $\mathbf{x}$ sobre la que U es constante. Tal superficie se llama equipotencial.

9.2 Ejemplo. En un punto $\mathbf{x}$ la fuerza que actúa sobre una partícula de masa m debida al campo gravitacional terrestre, es $\mathbf{F}(\mathbf{x}) = -m(0, 0, g)$. Demuéstrese que este campo de fuerzas es conservativo.

SOLUCIÓN. Deseamos demostrar que hay una función potencial U tal que $-\mathbf{D}U = \mathbf{F}$. Es este el caso si y sólo si

$$D_1 U = 0, \qquad D_2 U = 0, \qquad D_3 U = mg.$$

Una solución de estas ecuaciones es $U(x, y, z) = mgz$. Así pues, este campo de fuerza es conservativo. Las superficies equipotenciales son planos horizontales.

9.3 Ejemplo. Supongamos que una partícula de masa m con velocidad inicial $(a, 0, b)$ y posición inicial $(0, 0, 0)$ se mueve bajo la influencia del campo gravitacional de fuerzas $\mathbf{F}(x, y, z) = -m(0, 0, g)$. Verifíquese en este caso la ley de conservación de la energía.

SOLUCIÓN. La partícula se mueve de acuerdo a la ley de Newton $\mathbf{F} = m\mathbf{a}$. Así pues, si $\mathbf{x}(t)$ denota la posición de la partícula en el tiempo t,

$$\begin{aligned} \mathbf{x}(t) &= (0, 0, -g) \\ \dot{\mathbf{x}}(t) &= (a, 0, -gt+b) \\ \mathbf{x}(t) &= (at, 0, -\tfrac{1}{2}gt^2+bt). \end{aligned}$$

En cualquier tiempo t, como $U(x, y, z) = mgz$,

$$\begin{aligned} \tfrac{1}{2}m|\mathbf{v}(t)|^2 + U(\mathbf{x}(t)) &= \tfrac{1}{2}m(a^2+g^2t^2-2bgt+b^2)+mg(-\tfrac{1}{2}gt^2+bt) \\ &= \tfrac{1}{2}m(a^2+b^2). \end{aligned}$$

En un campo de fuerzas conservativo una partícula está en equilibrio estable en los puntos donde la energía potencial tiene un mínimo relativo.

9.4 Ejemplo. En el campo gravitacional de fuerzas $\mathbf{F}(x, y, z) = -m(0, 0, g)$ determínense los puntos sobre la superficie $9x^2+4y^2-yz+4 = 0$ donde una partícula de masa m estaría en equilibrio estable.

SOLUCIÓN. Como la función potencial es $U(x, y, z) = mgz$ donde m y g son positivos, determinaremos dónde z tiene un mínimo relativo. Resolviendo

la ecuación de la superficie para z, vemos que lo que hemos de encontrar son los puntos en que

$$f(x, y) = 4y + \frac{1}{y}(9x^2+4)$$

tiene un mínimo relativo. Igualando a cero la derivada de f, tenemos

$$D_1 f(x, y) = \frac{18x}{y} = 0$$

$$D_2 f(x, y) = 4 - \frac{1}{y^2}(9x^2+4) = 0$$

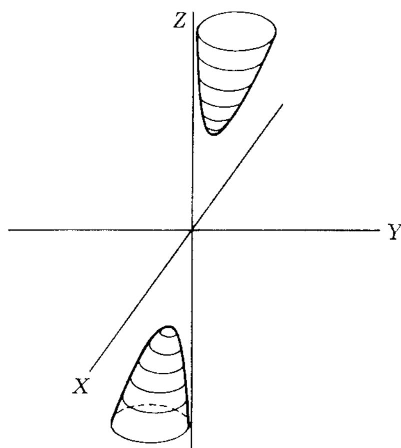


FIGURA 11

y, por tanto, $x = 0$ y $y = \pm 1$. Como

$$D_{1,1} f(x, y) = \frac{18}{y}, \quad D_{1,2} f(x, y) = -\frac{18x}{y^2}, \quad D_{2,2} f(x, y) = \frac{2}{y^3}(9x^2+4),$$

la expresión $D_{1,1} f D_{2,2} f - (D_{1,2} f)^2$ es positiva en los puntos $(0, 1)$ y $(0, -1)$. Además, $D_{1,1} f(0, 1) > 0$ y $D_{1,1} f(0, -1) < 0$. Por tanto, f tiene un mínimo relativo en $(0, 1)$ y un máximo relativo en $(0, -1)$. Por tanto, el único punto de equilibrio estable sobre la superficie dada es el $(0, 1, 8)$. La gráfica de la superficie aparece bosquejada en la figura 11.

Problemas

1. El campo de fuerzas de la atracción newtoniana está definido por $\mathbf{F}(\mathbf{x}) = -\dfrac{k}{|\mathbf{x}|^3}\mathbf{x}$, donde $k > 0$.

a) Demuéstrese que este campo de fuerza es conservativo.
b) ¿Cuáles son las superficies equipotenciales?
c) Verifíquese que en el punto $(2, 1, 2)$ la fuerza es ortogonal a la superficie equipotencial que pasa por $(2, 1, 2)$.
d) Encuéntrese la máxima razón de cambio del potencial en el punto $(2, 1, 2)$.
e) Si una partícula sólo puede moverse sobre la cúbica alabeada $\mathbf{x}(t) = (t, t^2+3, t^3-2)$, determínense los puntos de equilibrio estable.

2. Si $\mathbf{F}$ es un campo de fuerzas conservativo definido sobre un conjunto abierto y conexo, demuéstrese que cualesquier dos funciones potenciales de $\mathbf{F}$ pueden diferir solamente en una constante.

3. Supongamos que la fuerza que actúa sobre una partícula de masa m en el punto $\mathbf{x}$ es $\mathbf{F}(\mathbf{x}) = -mk\mathbf{x}$, donde $k > 0$.

a) Demuéstrese que $\mathbf{F}$ es conservativo.
b) Verifíquese la ley de conservación de la energía para este campo de fuerzas.
c) ¿Cuáles son las superficies equipotenciales?
d) Encuéntrese el trabajo hecho por $\mathbf{F}$ cuando la partícula se mueve desde el punto $(1, 0, 3)$ hasta el punto $(-2, 1, 5)$.
e) Si la partícula sólo puede moverse sobre el elipsoide $x^2+3y^2+4z^2 = 12$, encuéntrense los puntos de equilibrio estable.

4. Supongamos que la fuerza sobre una partícula en el punto (x, y, z) es

$$\mathbf{F}(x, y, z) = (y, -x, 0).$$

a) ¿Es conservativo este campo de fuerzas?
b) Encuéntrese el trabajo hecho al mover una partícula desde el punto $(1, 0, 0)$ hasta el $(-1, 0, 0)$ a lo largo de la mitad superior de la circunferencia unitaria en el plano XY.
c) Encuéntrese el trabajo hecho al mover una partícula desde $(1, 0, 0)$ a $(-1, 0, 0)$ a lo largo del eje X.

5. Si una partícula de masa m sujeta al campo gravitacional terrestre sólo puede moverse sobre la elipse

$$(2x+y+z)^2+y^2 = 1, \qquad 2x-3y = 0,$$

determínense los puntos de equilibrio estable.

10. RESUMEN

Estudiamos el cálculo diferencial de funciones vectoriales de una variable real en el capítulo 3 y el de funciones reales de un vector en el capítulo 4. En este capítulo completamos nuestra discusión del cálculo diferencial considerando el caso general de funciones vectoriales de un vector. Como los fundamentos de este caso general se siguen fácilmente del material estudiado en los capítulos 3 y 4, en este capítulo no hemos tenido realmente mucho nuevo que aprender.

Después de discutir los fundamentos del cálculo diferencial, consideramos algunas aplicaciones de las funciones vectoriales de un vector. La superficie se definió como una función de R^2 en R^3. Con el fin de discutir la aplicación física a los campos de fuerzas y al trabajo, introdujimos la integral curvilínea de una función vectorial a lo largo de una curva. Es esta esencialmente una integral unidimensional. En el próximo capítulo consideraremos integrales múltiples; es decir, integrales de funciones reales sobre conjuntos de dimensión mayor que uno.

Problemas de repaso

1. Determínense $D\mathbf{f}$ en $\mathbf{x}$ cuando

a) $\mathbf{f}(x, y, z) = (\cos xz,\ y^2 z)$, $\mathbf{x} = (2, 1, 1)$

b) $\mathbf{f}(x, y) = (\ln|x+y|,\ 2xy-3)$, $\mathbf{x} = (3, -5)$.

2. Encuéntrense todas las derivadas parciales de primero y segundo orden de $\mathbf{f}$ cuando

a) $\mathbf{f}(x, y, z) = \left(z \tan \dfrac{x}{y},\ xyz\right)$

b) $\mathbf{f}(x, y) = (\sqrt{x+y},\ e^{x+y})$.

3. Si u, v, x y y están relacionadas por las ecuaciones

$$\begin{aligned} xu^2 - y^2 v &= -2 \\ 2y^2 u + xyv &= 5 \end{aligned}$$

encuéntrense $\dfrac{\partial u}{\partial x}$, $\dfrac{\partial u}{\partial y}$, $\dfrac{\partial v}{\partial x}$ y $\dfrac{\partial v}{\partial y}$.

4. En la hipótesis de que f es diferenciable, exprésese

$$\frac{\partial}{\partial x} f(xy - x,\ x^2 + y)$$

en términos de las derivadas parciales de f.

5. En la hipótesis de que g es diferenciable, si $u = g(x, y)$ y $x = r\cos\theta$, $y = r$ senh θ, demuéstrese que

$$\left(\frac{\partial u}{\partial x}\right)^2 - \left(\frac{\partial u}{\partial y}\right)^2 = \left(\frac{\partial u}{\partial r}\right)^2 - \frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2$$

6. Determínese el plano tangente al paraboloide elíptico

$$x = v\cos u$$
$$y = 3v \operatorname{sen} u$$
$$z = v^2, \qquad u\in[0, 2\pi],\ v\in[0, \infty\rangle$$

en el punto $(-2, 0, 4)$.

7. Dada la elipse $3x^2+2xy+y^2 = 1$ con centro en el origen, encuéntrense sus puntos más lejanos del origen, determinando así su eje mayor.

8. Evalúese $\int_{\mathscr{C}} y\,dx + 2x\,dy$ cuando

a) $\mathscr{C}$ es la frontera del rectángulo con vértices $(3, -2)$, $(3, 2)$, $(-3, 2)$ y $(-3, -2)$ recorrida en dirección levógira.
b) $\mathscr{C}$ es la elipse $x = 3\cos t$, $y = 2$ sen t, $t\in[0, 2\pi]$.

9. Sea $\mathbf{F}$ un campo de fuerzas definido por $\mathbf{F}(\mathbf{x}) = -\dfrac{k}{|\mathbf{x}|^2}\mathbf{x}$, donde $k > 0$.

a) Demuéstrese que $\mathbf{F}$ es conservativo.
b) ¿Cuáles son las superficies equipotenciales?
c) Verifíquese que en el punto $(3, 2, 5)$ la fuerza es ortogonal a la superficie equipotencial que contiene a $(3, 2, 5)$.
d) Encuéntrese la razón máxima de cambio del potencial en el punto $(3, 2, 5)$.
e) Encuéntrese el trabajo hecho al mover una partícula desde el punto $(3, 2, 5)$ hasta el $(1, 0, -4)$.
f) Si una partícula está reducida a moverse sobre el elipsoide $\dfrac{x^2}{4} + y^2 + \dfrac{z^2}{9} = 1$, encuéntrense puntos cualesquiera en los que la partícula estaría en equilibrio estable.





Integrales múltiples

1. INTRODUCCIÓN

En la introducción al cálculo hemos estudiado la integral (definida) de Riemann, $\int_a^b f = \int_a^b f(x)\,dx$, de una función real de variable real. En este capítulo consideraremos la generalización del concepto de integral a dimensiones más altas —para funciones reales de diversas variables reales— y a conjuntos que son más generales que intervalos. Comenzaremos primero con la extensión en dimensión y posteriormente consideraremos integrales sobre conjuntos cerrados y acotados. Aunque la extensión a dimensiones más altas aumenta el problema del cálculo de la evaluación de la integral, la generalización es directa y natural. La extensión es tanto de interés matemático como de importancia considerable en las aplicaciones. Comen-

zaremos con la dimensión dos, donde el cuadro geométrico es completamente claro. La generalización a dimensiones arbitrarias es luego un fácil paso.

2. INTEGRALES DOBLES

Definiremos primero la integral definida para intervalos en R^2. En anteriores capítulos encontramos conveniente tomar el interior de una esfera n-dimensional como nuestra generalización de un intervalo unidimensional. Nos estábamos ocupando de vectores y la distancia entre puntos y el interior de una esfera es la generalización natural del intervalo en este caso. El interior de una esfera nos proporciona una generalización de la vecindad unidimensional que es expresable en términos de distancia de un modo sencillo: $\mathcal{N}(\mathbf{x}_0; \delta) = \{\mathbf{x} \mid |\mathbf{x}-\mathbf{x}_0| < \delta\}$. Aquí nos ocuparemos de los componentes de los vectores y consideraremos nuestro espacio como un producto cartesiano de espacios unidimensionales. Para nuestros propósitos presentes es más conveniente y natural convenir en que nuestros intervalos son rectángulos.

2.1 Definición. *Si* $\mathbf{a} = (a_1, a_2)$ *y* $\mathbf{b} = (b_1, b_2)$ *con* $a_1 \leqslant b_1$, $a_2 \leqslant b_2$, *entonces el* ***intervalo cerrado*** $[\mathbf{a}, \mathbf{b}]$ *en* R^2 *es el conjunto de todos los puntos* $\mathbf{x} = (x, y) \in R^2$ *tales que*

$$a_1 \leqslant x \leqslant b_1 \quad y \quad a_2 \leqslant y \leqslant b_2.$$

El intervalo cerrado $[\mathbf{a}, \mathbf{b}]$ en R^2 es un rectángulo con vértices en los puntos (a_1, a_2), (b_1, a_2), (b_1, b_2) y (a_1, b_2). Los lados del rectángulo son paralelos a los ejes coordenados (figura 1). Las longitudes de los lados de $[\mathbf{a}, \mathbf{b}]$ son los números b_1-a_1 y b_2-a_2. Si $b_1-a_1 = b_2-a_2$, entonces $[\mathbf{a}, \mathbf{b}]$ es un cuadrado.

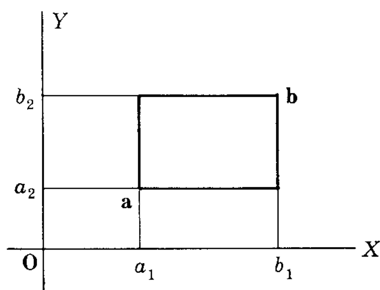


FIGURA 1

2.2 Definición. *El área del intervalo* $[\mathbf{a}, \mathbf{b}]$, *denotada por* $A([\mathbf{a}, \mathbf{b}])$, *es,*

por definición, el producto de las longitudes de los lados, es decir,

$$A([\mathbf{a}, \mathbf{b}]) = (b_1 - a_1)(b_2 - a_2).$$

Si $[\mathbf{a}, \mathbf{b}]$ es un intervalo cerrado en R, definimos una partición P de $[a, b]$ como un conjunto finito de números $x_0, x_1, \ldots, x_k$ tales que $a = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_{j-1} \leqslant x_j \leqslant \ldots \leqslant x_k = b$. Denotamos una partición de $[a, b]$ por $P = \{x_j \mid j = 0, 1, \ldots, k\}$, y definimos la *norma* o *malla* de P, lo que escribimos $|P|$, por

$$|P| = \text{máx}\,\{x_j - x_{j-1} \mid j = 1, \ldots, k\}.$$

Generalizaremos estos conceptos para intervalos en R^2.

2.3 Definición. *Sea* $[\mathbf{a}, \mathbf{b}] \subset R^2$ *con* $\mathbf{a} = (a_1, a_2)$ *y* $\mathbf{b} = (b_1, b_2)$. *Si* $P_1 = \{x_{j_1} \mid j_1 = 0, 1, \ldots, k_1\}$ *es una partición de* $[a_1, b_1]$ *y* $P_2 = \{y_{j_2} \mid j_2 = 0, 1, \ldots, k_2\}$ *es una partición de* $[a_2, b_2]$, *entonces*

$$P = P_1 \times P_2 = \{(x_{j_1}, y_{j_2}) \mid j_1 = 0, 1, \ldots, k_1 ;\ j_2 = 0, 1, \ldots, k_2\}$$

se dice que es una ***partición*** *de* $[\mathbf{a}, \mathbf{b}]$ *y definimos la* ***norma*** *de* P, *escrito* $|P|$, *por*

$$|P| = \text{máx}\,\{|P_1|, |P_2|\}.$$

Así pues, una partición P subdivide al rectángulo $[\mathbf{a}, \mathbf{b}]$ (figura 2) en cierto número de rectángulos más pequeños. La norma de la partición es la mayor de las dimensiones (largo o ancho) de todos estos rectángulos más pequeños y mide la finura de la partición. Si la partición P_1 subdivide a $[a_1, b_1]$ en k_1 subintervalos y P_2 subdivide a $[a_2, b_2]$ en k_2 subintervalos, entonces $P = P_1 \times P_2$ subdivide a $[\mathbf{a}, \mathbf{b}]$ en $k = k_1 k_2$ subintervalos. Los subintervalos bidimensionales de $[\mathbf{a}, \mathbf{b}]$ obtenidos por una partición P de $[\mathbf{a}, \mathbf{b}]$ pueden enumerarse consecutivamente y denotarse por $\mathscr{R}_i$ con $i = 1, \ldots, k$ donde $k = k_1 k_2$. La figura 2 ilustra una partición de un intervalo $[\mathbf{a}, \mathbf{b}]$ de R^2 con $k_1 = 5$ y $k_2 = 4$.

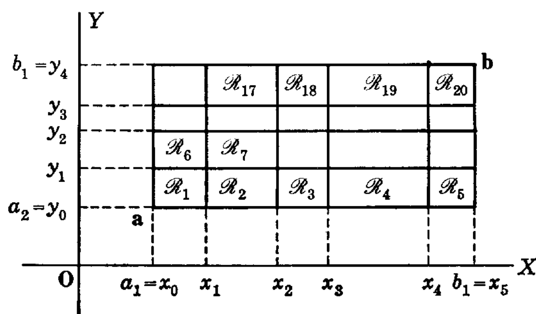


FIGURA 2

Sea f una función real que está acotada sobre $[\mathbf{a}, \mathbf{b}]$ de modo que hay números m y M tales que $m \leqslant f(\mathbf{x}) \leqslant M$ para todo $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$. Definamos

2.4
$$m_i(f) = \text{ínf}\,\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\}$$
$$M_i(f) = \sup\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\}\,.$$

El supremo de un conjunto $\mathscr{S}$ de números reales, denotado por sup $\mathscr{S}$ (o lub $\mathscr{S}$), es un número c tal que para todo $x \in \mathscr{S}$, $x \leqslant c$ y si $b < c$ entonces hay un $x \in \mathscr{S}$ tal que $x > b$; es decir, el supremo de $\mathscr{S}$ es una cota superior de $\mathscr{S}$ y cualquier número menor que él no es una cota superior de $\mathscr{S}$. El ínfimo de $\mathscr{S}$, ínf $\mathscr{S}$ (o glb $\mathscr{S}$), se define de un modo análogo. Es una propiedad básica del sistema de los números reales que todo conjunto no vacío $\mathscr{S}$ de números reales tiene un supremo si $\mathscr{S}$ está superiormente acotado, y tiene un ínfimo si $\mathscr{S}$ está inferiormente acotado.

Como estamos suponiendo que f está acotado sobre $[\mathbf{a}, \mathbf{b}]$, $M_i(f)$ y $m_i(f)$ existen para todo $i = 1, \ldots, k$, y

2.5
$$m \leqslant m(f) \leqslant M_i(f) \leqslant M.$$

La definición de una integral la daremos en términos de sumas de los siguientes tipos:

$$L(f, P) = \sum_{i=1}^{k} m_i(f)\,A(\mathscr{R}_i)$$

y

$$U(f, P) = \sum_{i=1}^{k} M_i(f)\,A(\mathscr{R}_i)$$

donde $A(\mathscr{R}_i)$ es el área del i-ésimo subintervalo $\mathscr{R}_i$ en la partición P. Llamamos a $L(f, P)$ "suma inferior", y a $U(f, P)$ "suma superior" correspondientes a la partición P.

Estas sumas tienen una sencilla interpretación geométrica para funciones no negativas en un intervalo $[\mathbf{a}, \mathbf{b}]$ en R^2, aunque debemos recordar que la sola restricción que hemos impuesto sobre nuestras funciones es que han de ser acotadas. La gráfica de una función no negativa y acotada sobre un intervalo $[\mathbf{a}, \mathbf{b}]$ aparece representada en la figura 3. La suma inferior

$$L(f, P) = m_1(f)\,A(\mathscr{R}_1) + \ldots + m_k(f)\,A(\mathscr{R}_k) = \sum_{i=1}^{k} m_i(f)\,A(\mathscr{R}_i)$$

es la suma de los volúmenes de los paralelepípedos interiores. La suma superior

$$U(f, P) = M_1(f)\,A(\mathscr{R}_1) + \ldots + M_k(f)\,A(\mathscr{R}_k) = \sum_{i=1}^{k} M_i(f)\,A(\mathscr{R}_i)$$

es la suma de los volúmenes de los paralelepípedos exteriores.

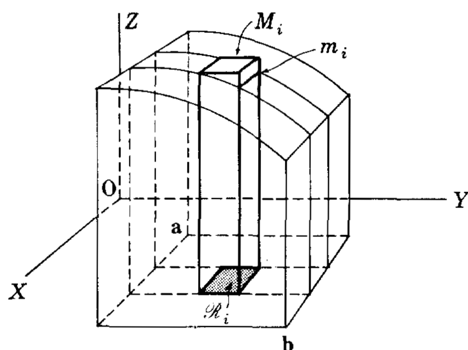


FIGURA 3

De acuerdo con 2.5, si f está acotada sobre $[\mathbf{a}, \mathbf{b}]$ y P es una partición de $[\mathbf{a}, \mathbf{b}]$, entonces

$$mA([\mathbf{a},\mathbf{b}]) = \sum_{i=1}^{k} mA(\mathscr{R}_i) \leqslant \sum_{i=1}^{k} m_i(f)A(\mathscr{R}_i)$$

$$\leqslant \sum_{i=1}^{k} M_i(f)A(\mathscr{R}_i) \leqslant \sum_{i=1}^{k} MA(\mathscr{R}_i) = MA([\mathbf{a},\mathbf{b}])$$

o bien

2.6 $$mA([\mathbf{a}, \mathbf{b}]) \leqslant L(f, P) \leqslant U(f, P) \leqslant MA([\mathbf{a}, \mathbf{b}]).$$

Sea $\mathscr{P}$ el conjunto de todas las particiones del intervalo $[\mathbf{a}, \mathbf{b}]$. La desigualdad 2.6 se verifica para cada partición P en $\mathscr{P}$ y nos muestra que el conjunto de todos los números $\{L(f, P) \mid P \in \mathscr{P}\}$ —el conjunto de todas las sumas inferiores obtenidas tomando todas las posibles particiones de $[\mathbf{a}, \mathbf{b}]$— tiene una cota superior; a saber, $MA([\mathbf{a}, \mathbf{b}])$. El conjunto $\{L(f, P) \mid P \in \mathscr{P}\}$ tiene, por tanto, un supremo. Análogamente, el conjunto $\{U(f, P) \mid P \in \mathscr{P}\}$ tiene una cota inferior $mA([\mathbf{a}, \mathbf{b}])$ y, por tanto, $\{U(f, P) \mid P \in \mathscr{P}\}$ tiene un ínfimo. Este supremo y este ínfimo son suficientemente importantes para que introduzcamos nombres y símbolos que los denominen y denoten.

2.7 Definición. *Definimos*

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \inf \{L(f, P) \mid P \in \mathscr{P}\}$$

y

$$\overline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \sup \{U(f, P) \mid P \in \mathscr{P}\}.$$

$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f$ *se llama* ***integral inferior*** *de* f *sobre* $[\mathbf{a}, \mathbf{b}]$, *y* $\overline{\int_{\mathbf{a}}^{\mathbf{b}}} f$ *se llama* ***integral superior*** *de* f *sobre* $[\mathbf{a}, \mathbf{b}]$.

La discusión que precede a esta definición establece la existencia de $\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f$ e $\overline{\int_{\mathbf{a}}^{\mathbf{b}}} f$ para todas las funciones f que son acotadas en $[\mathbf{a}, \mathbf{b}]$.

Si P y P' son particiones de un intervalo $[\mathbf{a}, \mathbf{b}]$, $P \subset P'$ significa que cada punto de división de P es también un punto de división de P'. Cuando éste es el caso, P' se dice que es un *refinamiento* de P. Demostraremos que ningún refinamiento de una partición hace decrecer la suma inferior ni aumentar la suma superior. Enunciado en forma precisa, probaremos que para toda función acotada f:

2.8 Lema. *Si* $P \subset P'$, *entonces* $L(f, P) \leqslant L(f, P')$ *y* $U(f, P') \leqslant U(f, P)$.

PRUEBA. Si $P = P'$ el lema es obviamente cierto. Supongamos que $P = P_1 \times P_2$ y $P' = P_1' \times P_2'$, con $P \neq P'$ y $P \subset P'$. Supongamos también que $P_1 \neq P_1'$. Sea x_j' el primer punto de división de P_1' que no está en P_1. Entonces, para algún l, $x_{l-1} < x_j' < x_l$. Definamos $P_1{}^1$ por

$$P_1{}^1 = \{x_0, x_1, \ldots, x_{l-1}, x_j', x_l, \ldots, x_{k_1}\}$$

y

$$P^1 = P_1{}^1 \times P_2.$$

Las subregiones $\mathscr{R}_i$ limitadas por x_{l-1} y x_l están, cada una, divididas en dos partes distintas por la inserción de x_j'. Un caso particular de esto aparece ilustrado en la figura 4.

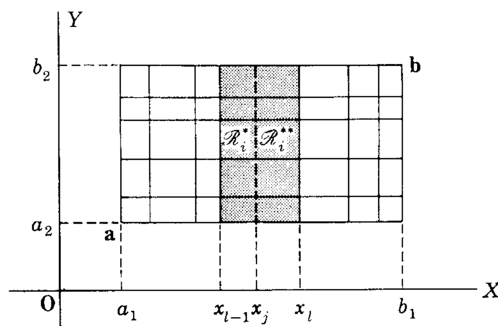


FIGURA 4

Sea $\mathscr{R}_i$ una de las regiones que está dividida por la inserción de x_j' y sean $\mathscr{R}_i{}^*$ y $\mathscr{R}_i{}^{**}$ las dos subregiones en que la $\mathscr{R}_i$ se ha dividido. Definamos

$$m_i{}^*(f) = \inf\,\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i{}^*\}$$

y

$$m_i^{**}(f) = \text{ínf}\,\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i^{**}\}.$$

De la definición de $m_i(f)$ se deduce que $m_i(f) \leqslant m_i^*(f)$ y $m_i(f) \leqslant m_i^{**}(f)$. Por tanto

$$\begin{aligned} m_i(f)\,A(\mathscr{R}_i) &= m_i(f)\,A(\mathscr{R}_i^*) + m_i(f)\,A(\mathscr{R}_i^{**}) \\ &\leqslant m_i^*(f)\,A(\mathscr{R}_i^*) + m_i^{**}(f)\,A(\mathscr{R}_i^{**}) \end{aligned}$$

de donde

$$\begin{aligned} L(f, P) &= m_1(f)\,A(\mathscr{R}_1) + \ldots + m_i(f)\,A(\mathscr{R}_i) + \ldots + m_k(f)\,A(\mathscr{R}_k) \\ &\leqslant m_1(f)\,A(\mathscr{R}_1) + \ldots + m_i^*(f)\,A(\mathscr{R}_i^*) \\ &\qquad + m_i^{**}(f)\,A(\mathscr{R}_i^{**}) + \ldots + m_k(f)\,A(\mathscr{R}_k). \end{aligned}$$

Hay un número finito de subregiones $\mathscr{R}_i$ que están subdivididas por la adición de x_j'. Luego, por repetición del anterior argumento un número finito de veces, tenemos

$$L(f, P) \leqslant L(f, P^1).$$

Repitiendo todo el proceso un número finito de veces, podemos añadir todos los puntos de P_1' que no están en P_1 y en la misma forma podemos añadir los puntos de P_2' que no están en P_2. Así obtenemos $L(f, P) \leqslant L(f, P')$. De una forma análoga —con todas las desigualdades que aquí aparecieron invertidas— obtenemos que $U(f, P) \geqslant U(f, P')$.

Aplicamos ahora el lema 2.8 junto con 2.6 para obtener la siguiente importante propiedad sobre integrales superiores e inferiores.

2.9 Lema. *Si f es acotada sobre* $[\mathbf{a}, \mathbf{b}]$, *entonces, para cualquier partición P de* $[\mathbf{a}, \mathbf{b}]$,

$$L(f, P) \leqslant \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant U(f, P).$$

PRUEBA. Como $\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \sup\,\{L(f, P) \mid P \in \mathscr{P}$, se sigue que $L(f, P) \leqslant \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f$. Análogamente, $\overline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant U(f, P)$. Queda, pues, por mostrar que $\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f$.

Sea $P = P_1 \times P_2$ y $Q = Q_1 \times Q_2$ un par cualquiera de particiones de $[\mathbf{a}, \mathbf{b}]$ y sea $P' = (P_1 \cup Q_1) \times (P_2 \cup Q_2)$. Es claro que $P \subset P'$ y $Q \subset P'$, es decir, que P' es un refinamiento tanto de P como de Q. Luego, de acuerdo con el lema 2.8,

$$L(f, P) \leqslant L(f, P') \quad \text{y} \quad U(f, P') \leqslant U(f, Q).$$

Como, según 2.6, $L(f, P') \leqslant U(f, P')$, tenemos

$$L(f, P) \leqslant U(f, Q)$$

para cualquier par de particiones P y Q de $[\mathbf{a}, \mathbf{b}]$. Así pues, para toda partición Q, $U(f, Q)$ es una cota superior de $\{L(f, P) \mid P \in \mathscr{P}\}$. Como $\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f$ es el supremo de $\{L(f, P) \mid P \in \mathscr{P}\}$,

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant U(f, Q) \text{ para todo } Q \in \mathscr{P},$$

e $\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f$ es una cota superior de $\{U(f, Q) \mid Q \in \mathscr{P}\}$. Como $\overline{\int_{\mathbf{a}}^{\mathbf{b}}} f$ es el supremo de $\{U(f, Q) \mid Q \in \mathscr{P}\}$,

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f.$$

Esto completa la prueba.

Definimos ahora las funciones integrables y la integral definida de una función integrable sobre un intervalo.

2.10 Definición. *Una función f sobre $[\mathbf{a}, \mathbf{b}]$ se dice que es **(de Riemann) integrable sobre** $[\mathbf{a}, \mathbf{b}]$ si f es acotada sobre $[\mathbf{a}, \mathbf{b}]$ e*

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f.$$

*Si f es integrable sobre $[\mathbf{a}, \mathbf{b}]$ entonces la **integral definida (de Riemann) de f sobre** $[\mathbf{a}, \mathbf{b}]$, escrita $\int_{\mathbf{a}}^{\mathbf{b}} f$, está definida por*

$$\int_{\mathbf{a}}^{\mathbf{b}} f = \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f.$$

Puede también usarse para denotar la integral de f la notación $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\, d\mathbf{x}$. La integral de una función f sobre un intervalo $[\mathbf{a}, \mathbf{b}] \subset \mathrm{R}^2$ se llama *integral doble*. El término "doble" se refiere a la dimensión del intervalo $[\mathbf{a}, \mathbf{b}]$. Las notaciones

$$\iint_{\mathbf{a}}^{\mathbf{b}} f(x, y)\, dx\, dy \quad \text{e} \quad \iint_{\mathbf{a}}^{\mathbf{b}} f(x, y)\, dA$$

se usan a veces para denotar la integral doble de f sobre $[\mathbf{a}, \mathbf{b}]$.

El siguiente resultado es una consecuencia directa del lema 2.9.

2.11 Teorema. *Si f es integrable sobre* $[\mathbf{a}, \mathbf{b}] \subset \mathrm{R}^2$ y *P es una partición cualquiera de* $[\mathbf{a}, \mathbf{b}]$, *entonces*

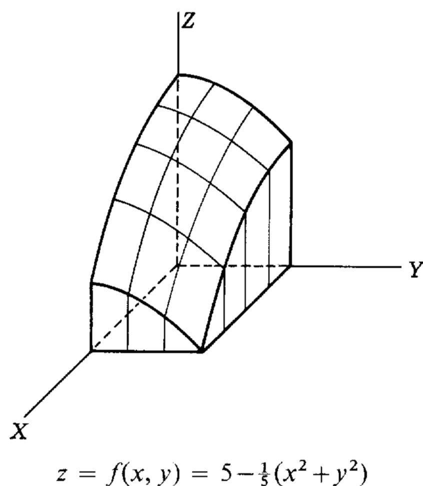
$$L(f, P) \leqslant \int_{\mathbf{a}}^{\mathbf{b}} f \leqslant U(f, P).$$

PRUEBA. De acuerdo con el lema 2.9, para cualquier partición P de $[\mathbf{a}, \mathbf{b}]$,

$$L(f, P) \leqslant \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant U(f, P).$$

Como f es integrable sobre $[\mathbf{a}, \mathbf{b}]$, $\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f$ y el teorema sigue.

Este teorema muestra que cualquier suma inferior es una cota inferior para la integral definida y cualquier suma superior es una cota superior. Ilustraremos el método para la aproximación al valor de la integral mediante el cálculo de las sumas superiores e inferiores.



$z = f(x, y) = 5 - \frac{1}{5}(x^2 + y^2)$

FIGURA 5

2.12 Ejemplo. Supongamos que $f(x, y) = 5 - \frac{1}{5}(x^2 + y^2)$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$ donde $\mathbf{a} = (0, 0)$ y $\mathbf{b} = (4, 3)$ (lo que más adelante sabremos que es cierto). Calcúlese en forma aproximada $\int_{\mathbf{a}}^{\mathbf{b}} f$.

SOLUCIÓN. Sea $P = P_1 \times P_2$ donde $P_1 = \{0, 1, 2, 3, 4\}$ y $P_2 = \{0, 1, 2, 3\}$. Esto está ilustrado en la figura 5. Como la función es monótona decreciente

tanto respecto a x como respecto a y, y $A(\mathscr{R}_i) = 1$ para todo i, tenemos

$$\begin{aligned} L(f, P) &= \sum_{i=1}^{12} m_i(f)A(\mathscr{R}_i) = f(1,1)+f(2,1)+f(3,1)+f(4,1)+f(1,2) \\ &\quad +f(2,2)+f(3,2)+f(4,2)+f(1,3)+f(2,3)+f(3,3)+f(4,3) \\ &= \tfrac{23}{5}+\tfrac{20}{5}+\tfrac{15}{5}+\tfrac{8}{5}+\tfrac{20}{5}+\tfrac{17}{5}+\tfrac{12}{5}+\tfrac{5}{5}+\tfrac{15}{5}+\tfrac{12}{5}+\tfrac{7}{5}+0 \\ &= 30.8 \end{aligned}$$

y

$$\begin{aligned} U(f,P) &= \sum_{i=1}^{12} M_i(f)A(\mathscr{R}_i) = f(0,0)+f(1,0)+f(2,0)+f(3,0)+f(0,1) \\ &\quad +f(1,1)+f(2,1)+f(3,1)+f(0,2)+f(1,2)+f(2,2)+f(3,2) \\ &= 47.6. \end{aligned}$$

Entonces aproximamos

$$\iint_{(0,0)}^{(4,3)} [5-\tfrac{1}{5}(x^2+y^2)]\,dx\,dy \quad \text{por} \quad \tfrac{1}{2}[L(f, P)+U(f, P)] = 39.2.$$

Como $L(f, P) \leqslant \int_{\mathbf{a}}^{\mathbf{b}} f \leqslant U(f, P)$ implica

$$\tfrac{1}{2}[L(f,P)-U(f,P)] \leqslant \int_{\mathbf{a}}^{\mathbf{b}} f-\tfrac{1}{2}[L(f,P)+U(f,P)] \leqslant \tfrac{1}{2}[U(f,P)-L(f, P)],$$

vemos que

2.13 $$\left| \int_{\mathbf{a}}^{\mathbf{b}} f-\tfrac{1}{2}[L(f, P)+U(f, P)] \right| \leqslant \tfrac{1}{2}[U(f, P)-L(f, P)].$$

El número $\frac{1}{2}[U(f, P)-L(f, P)]$ es una cota superior del error cometido al aproximar $\int_{\mathbf{a}}^{\mathbf{b}} f$ por $\frac{1}{2}[L(f, P)+U(f, P)]$. Así pues, en nuestro ejemplo el error al tomar como valor de la integral el de 39.2 es menor que o igual a $\frac{1}{2}(47.6-30.8) = 8.4$. La aproximación es —como el aspecto de la figura 5 nos sugiere— mucho mejor que esto. En realidad, el valor de la integral es 40.

A continuación demostraremos que una función acotada es integrable sobre el intervalo $[\mathbf{a}, \mathbf{b}]$ si y sólo si la diferencia entre una suma superior y la suma inferior correspondiente puede hacerse arbitrariamente pequeña.

2.14 Teorema. *Una función acotada f es integrable sobre $[\mathbf{a}, \mathbf{b}]$ si y sólo si para cada $\varepsilon > 0$ existe una partición P de $[\mathbf{a}, \mathbf{b}]$ con la propiedad de que $U(f, P)-L(f, P) < \varepsilon$.*

PRUEBA. Supongamos que para cada $\varepsilon > 0$ una tal P existe. De acuerdo con el lema 2.9,

$$0 \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f - \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant U(f, P) - L(f, P).$$

Como $U(f, P) - L(f, P) < \varepsilon$, se sigue que $\overline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f$, y f es integrable sobre $[\mathbf{a}, \mathbf{b}]$.

Recíprocamente, supongamos que f es integrable sobre $[\mathbf{a}, \mathbf{b}]$. Entonces $\int_{\mathbf{a}}^{\mathbf{b}} f = \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \int_{\mathbf{a}}^{\mathbf{b}} f$. Por tanto $\int_{\mathbf{a}}^{\mathbf{b}} f = \sup\{L(f, P)\} = \inf\{U(f, P)\}$. Luego para cada $\varepsilon > 0$ existe una partición P^1 tal que

$$\int_{\mathbf{a}}^{\mathbf{b}} f - L(f, P^1) < \frac{\varepsilon}{2}$$

y una partición P^2 tal que

$$U(f, P^2) - \int_{\mathbf{a}}^{\mathbf{b}} f < \frac{\varepsilon}{2}.$$

Añadiendo estas dos desigualdades obtenemos

$$U(f, P^2) - L(f, P^1) < \varepsilon.$$

Luego, para todo refinamiento común P de P^1 y P^2,

$$U(f, P) - L(f, P) \leqslant U(f, P^2) - L(f, P^1) < \varepsilon.$$

Y esto completa la prueba.

De acuerdo con el teorema 2.14, vemos que siempre es posible aproximar integrables definidas por medio de sus sumas superiores e inferiores tanto como deseemos.

Problemas

1. Encuéntrese el área del intervalo $[\mathbf{a}, \mathbf{b}]$ si

a) $\mathbf{a} = (1, 2)$, $\mathbf{b} = (3, 5)$ *b*) $\mathbf{a} = (0, 0)$, $\mathbf{b} = (3, 4)$
c) $\mathbf{a} = (-2, -1)$, $\mathbf{b} = (1, 3)$ *d*) $\mathbf{a} = (-3, 1)$, $\mathbf{b} = (2, 1)$.

2. Sean $\mathbf{a} = (0, 0)$, $\mathbf{b} = (1, 2)$, y $f(x, y) = 2x + y$. Encuéntrense $L(f, P)$ y $U(f, P)$ cuando $P = P_1 \times P_2$ si $P_1 = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ y $P_2 = \{0, \frac{1}{2}, 1, \frac{3}{2}, 2\}$.

3. Supongamos —como es el caso— que las funciones que a continuación aparecen son integrables sobre los intervalos dados. Demuéstrese que:

a) $\frac{1}{2} \leqslant \int_{\mathbf{a}}^{\mathbf{b}} f \leqslant \frac{3}{2}$, si $\mathbf{a} = (0, 0)$, $\mathbf{b} = (1, 1)$, $f(x, y) = x + y$

b) $1.5 \leqslant \int_{\mathbf{a}}^{\mathbf{b}} f \leqslant 4.5$, si $\mathbf{a} = (0, 0)$, $\mathbf{b} = \left(\frac{\pi}{2}, \frac{\pi}{2}\right)$, $f(x, y) = \operatorname{sen} x + \cos y$.

4. Para cada uno de los intervalos $\mathscr{R}_i$ de una cierta partición P de $[\mathbf{a}, \mathbf{b}]$ sea x_i un punto de $\mathscr{R}_i$. Demuéstrese que cualquiera que sea la selección de los x_i,

$$L(f, P) \leqslant \sum_{i=1}^{k} f(\mathbf{x}_i) A(\mathscr{R}_i) \leqslant U(f, P).$$

De ello se concluye que si f es integrable sobre $[\mathbf{a}, \mathbf{b}]$, entonces

$$\left| \int_{\mathbf{a}}^{\mathbf{b}} f - \sum_{i=1}^{k} f(\mathbf{x}_i) A(\mathscr{R}_i) \right| \leqslant U(f, P) - L(f, P).$$

3. PROPIEDADES BÁSICAS DE $\int_{\mathbf{a}}^{\mathbf{b}} f$

Estableceremos ahora algunas propiedades básicas de la integral doble $\int_{\mathbf{a}}^{\mathbf{b}} f$ que son análogas a las propiedades básicas de la integral unidimensional $\int_{a}^{b} f$.

3.1 *Si la función f es integrable sobre $[\mathbf{a}, \mathbf{b}]$ y c es una función constante de* R^2 *en* R, *entonces la función cf es integrable sobre $[\mathbf{a}, \mathbf{b}]$ e* $\int_{\mathbf{a}}^{\mathbf{b}} cf = c \int_{\mathbf{a}}^{\mathbf{b}} f$.

3.2 *Si c es una función constante de* R^2 *a* R, *entonces para todo intervalo* $[\mathbf{a}, \mathbf{b}]$ *en* R^2, *c es integrable sobre $[\mathbf{a}, \mathbf{b}]$ e* $\int_{\mathbf{a}}^{\mathbf{b}} c = cA([\mathbf{a}, \mathbf{b}])$.

3.3 *Si las funciones f y g son integrables sobre* $[\mathbf{a}, \mathbf{b}]$, *entonces la función $f+g$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$ e* $\int_{\mathbf{a}}^{\mathbf{b}} (f+g) = \int_{\mathbf{a}}^{\mathbf{b}} f + \int_{\mathbf{a}}^{\mathbf{b}} g$.

3.4 *La función f es integrable sobre* $[\mathbf{a}, \mathbf{b}]$ *si y sólo si las funciones f^+ y f^- definidas por las reglas*

$$f^+(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{si } f(\mathbf{x}) \geqslant 0 \\ 0 & \text{si } f(\mathbf{x}) < 0 \end{cases} \quad \text{y} \quad f^-(\mathbf{x}) = \begin{cases} 0 & \text{si } f(\mathbf{x}) \geqslant 0 \\ -f(\mathbf{x}) & \text{si } f(\mathbf{x}) < 0 \end{cases}$$

son ambas integrables sobre $[\mathbf{a}, \mathbf{b}]$.

3.5 *Si la función f es integrable sobre* $[\mathbf{a}, \mathbf{b}]$, *entonces f^2 es integrable sobre* $[\mathbf{a}, \mathbf{b}]$.

3.6 *Si las funciones f y g son integrables sobre* $[\mathbf{a}, \mathbf{b}]$ *entonces el producto fg es integrable sobre* $[\mathbf{a}, \mathbf{b}]$.

3.7 *Si las funciones f y g son integrables sobre* $[\mathbf{a}, \mathbf{b}]$ *y $f(\mathbf{x}) \leqslant g(\mathbf{x})$ para todo* $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$, *entonces* $\int_{\mathbf{a}}^{\mathbf{b}} f \leqslant \int_{\mathbf{a}}^{\mathbf{b}} g$.

3.8 *Si la función f es integrable sobre* $[\mathbf{a}, \mathbf{b}]$, *entonces f es integrable sobre* $[\mathbf{a}, \mathbf{b}]$ *e* $\left|\int_{\mathbf{a}}^{\mathbf{b}} f\right| \leqslant \int_{\mathbf{a}}^{\mathbf{b}} |f|$.

PRUEBA DE 3.1. Sea P una partición cualquiera de $[\mathbf{a}, \mathbf{b}]$ y denotemos por $\mathscr{R}_i$ el i-ésimo subintervalo de la partición P. Si $c \geqslant 0$, entonces para todo i tenemos

$$m_i(cf) = \inf\{cf(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\} = c \inf\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\} = cm_i(f)$$

de donde

$$L(cf, P) = \sum_{i=1}^{k} m_i(cf) A(\mathscr{R}_i) = \sum_{i=1}^{k} cm_i(f) A(\mathscr{R}_j) = cL(f, P).$$

De donde se sigue que

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} cf = c \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f.$$

Análogamente, para $c \geqslant 0$

$$\overline{\int_{\mathbf{a}}^{\mathbf{b}}} cf = c \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f.$$

Como f se supone integrable sobre $[\mathbf{a}, \mathbf{b}]$,

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} cf = c \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = c \int_{\mathbf{a}}^{\mathbf{b}} f = c \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} cf$$

y cf es también integrable sobre $[\mathbf{a}, \mathbf{b}]$ con

$$\int_{\mathbf{a}}^{\mathbf{b}} cf = c \int_{\mathbf{a}}^{\mathbf{b}} f.$$

Si $c < 0$, entonces

$$m_i(cf) = \inf\{cf(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\} = c \sup\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\} = cM_i(f)$$

y

$$M_i(cf) = \sup\{cf(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\} = c \inf\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\} = cm_i(f).$$

En este caso

$$\overline{\int_{\mathbf{a}}^{\mathbf{b}}} cf = c \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f \quad \text{e} \quad \underline{\int_{\mathbf{a}}^{\mathbf{b}}} cf = c \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f.$$

De nuevo, como f se supone integrable sobre $[\mathbf{a}, \mathbf{b}]$,

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} cf = c \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f = c \int_{\mathbf{a}}^{\mathbf{b}} f = c \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} cf$$

y en este caso cf es también integrable sobre $[\mathbf{a}, \mathbf{b}]$ con

$$\int_{\mathbf{a}}^{\mathbf{b}} cf = c \int_{\mathbf{a}}^{\mathbf{b}} f.$$

PRUEBA DE 3.2. Para cualquier partición P de $[\mathbf{a}, \mathbf{b}]$, $L(c, P) = cA([\mathbf{a}, \mathbf{b}]) = U(c, P)$.

De donde, según el lema 2.9,

$$cA([\mathbf{a}, \mathbf{b}]) \leqslant \underline{\int_{\mathbf{a}}^{\mathbf{b}}} c \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} c \leqslant cA([\mathbf{a}, \mathbf{b}])$$

y vemos que c es integrable sobre $[\mathbf{a}, \mathbf{b}]$ con

$$\int_{\mathbf{a}}^{\mathbf{b}} c = cA([\mathbf{a}, \mathbf{b}]).$$

Antes de probar 3.3 estableceremos el siguiente lema.

3.9 Lema. *Si f y g están acotados sobre* $[\mathbf{a}, \mathbf{b}]$, *entonces*

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} (f+g) \geqslant \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f + \underline{\int_{\mathbf{a}}^{\mathbf{b}}} g \quad \text{e} \quad \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f + \overline{\int_{\mathbf{a}}^{\mathbf{b}}} g \geqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} (f+g).$$

PRUEBA. Sea P una partición cualquiera de $[\mathbf{a}, \mathbf{b}]$ y denotemos por $\mathscr{R}_i$ el i-ésimo subintervalo de la partición. Para todo i tenemos

$$\begin{aligned} m_i(f+g) &= \inf \{f(\mathbf{x})+g(\mathbf{x}) \mid \mathbf{x}\in\mathscr{R}_i\} \\ &\geqslant \inf \{f(\mathbf{x}) \mid \mathbf{x}\in\mathscr{R}_i\} + \inf \{g(\mathbf{x}) \mid \mathbf{x}\in\mathscr{R}_i\} = m_i(f)+m_i(g) \end{aligned}$$

y

$$\begin{aligned} M_i(f+g) &= \sup \{f(\mathbf{x})+g(\mathbf{x}) \mid \mathbf{x}\in\mathscr{R}_i\} \\ &\leqslant \sup \{f(\mathbf{x}) \mid \mathbf{x}\in\mathscr{R}_i\} + \sup \{g(\mathbf{x}) \mid \mathbf{x}\in\mathscr{R}_i\} = M_i(f)+M_i(g). \end{aligned}$$

De estas desigualdades se sigue que si P^1 y P^2 son particiones de $[\mathbf{a}, \mathbf{b}]$

y P es un refinamiento común de P^1 y P^2, entonces

$$L(f, P^1)+L(g, P^2) \leqslant L(f, P)+L(g, P) \leqslant L(f+g, P) \leqslant \underline{\int_{\mathbf{a}}^{\mathbf{b}}} (f+g)$$

e

$$\overline{\int_{\mathbf{a}}^{\mathbf{b}}} (f+g) \leqslant U(f+g, P) \leqslant U(f, P)+U(g, P) \leqslant U(f, P^1)+U(g, P^2).$$

Como estas desigualdades se verifican para $P^1, P^2 \in \mathscr{P}$ arbitrarios, tenemos

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} (f+g) \geqslant \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f + \underline{\int_{\mathbf{a}}^{\mathbf{b}}} g \quad \text{e} \quad \overline{\int_{\mathbf{a}}^{\mathbf{b}}} (f+g) \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f + \overline{\int_{\mathbf{a}}^{\mathbf{b}}} g .$$

PRUEBA DE 3.3. De acuerdo con el lema 3.9 tenemos

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f + \underline{\int_{\mathbf{a}}^{\mathbf{b}}} g \leqslant \underline{\int_{\mathbf{a}}^{\mathbf{b}}} (f+g) \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} (f+g) \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f + \overline{\int_{\mathbf{a}}^{\mathbf{b}}} g .$$

Como f y g se suponen integrables sobre $[\mathbf{a}, \mathbf{b}]$, en esta relación debe verificarse la igualdad y $f+g$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$ con

$$\int_{\mathbf{a}}^{\mathbf{b}} (f+g) = \int_{\mathbf{a}}^{\mathbf{b}} f + \int_{\mathbf{a}}^{\mathbf{b}} g .$$

PRUEBA DE 3.4. Probaremos que

3.10 $$U(f, P)-L(f, P) = [U(f^+, P)-L(f^+, P)]+[U(f^-, P)-L(f^-, P)].$$

La propiedad 3.4 se sigue entonces de 3.10 por una aplicación del teorema 2.14.

Sea P una partición de $[\mathbf{a}, \mathbf{b}]$. Probaremos primero que sobre cada intervalo $\mathscr{R}_i$

3.11 $$M_i(f)-m_i(f) = M_i(f^+)-m_i(f^+)+M_i(f^-)-m_i(f^-).$$

Consideraremos tres casos:

Caso 1. $m_i(f) \geqslant 0$. Aquí

$$M_i(f^+) = M_i(f), \quad m_i(f^+) = m_i(f), \quad M_i(f^-) = m_i(f^-) = 0$$

y 3.11 se verifica.

Caso 2. $M_i(f) \leqslant 0$. Aquí

$$M_i(f^+) = m_i(f^+) = 0, \quad M_i(f^-) = -m_i(f), \quad m_i(f^-) = -M_i(f)$$

y 3.11 se verifica también en este caso.

Caso 3. $m_i(f) < 0 < M_i(f)$. En este caso

$$M_i(f^+) = M_i(f), \quad m_i(f^+) = 0, \quad M_i(f^-) = -m_i(f), \quad m_i(f^-) = 0$$

de modo que

$$M_i(f)-m_i(f) = M_i(f^+)+M_i(f^-) = M_i(f^+)-m_i(f^+)+M_i(f^-)-m_i(f^-)$$

y, de nuevo, 3.11 se verifica.

Vemos pues que en los tres casos se verifica 3.11. Si multiplicamos 3.11 por $A(\mathscr{R}_i)$ y sumamos sobre i, obtenemos 3.10.

Cada una de las diferencias $U(f, P)-L(f, P)$, $U(f^+, P)-L(f^+, P)$ y $U(f^-, P)-L(f^-, P)$ es no negativa. Si f es integrable sobre $[\mathbf{a}, \mathbf{b}]$, entonces según el teorema 2.14 para cada $\varepsilon > 0$, hay una partición P tal que $U(f, P)-L(f, P) < \varepsilon$. Luego, por 3.10, $U(f^+, P)-L(f^+, P) < \varepsilon$ y $U(f^-, P)-L(f^-, P) < \varepsilon$, y, de nuevo según el teorema 2.14, f^+ y f^- son integrables sobre $[\mathbf{a}, \mathbf{b}]$.

Recíprocamente, si f^+ y f^- son integrables sobre $[\mathbf{a}, \mathbf{b}]$, entonces, por el teorema 2.14 para cada $\varepsilon > 0$, hay particiones P^1 y P^2 tales que $U(f^+, P^1)-L(f^+, P^1) < \varepsilon/2$ y $U(f^-, P^2)-L(f^-, P^2) < \varepsilon/2$. Sea P un refinamiento común de P^1 y P^2. Entonces

$$\begin{aligned} U(f, P)-L(f, P) &= U(f^+, P)-L(f^+, P)+U(f^-, P)-L(f^-, P) \\ &\leqslant U(f^+, P^1)-L(f^+, P^1)+U(f^-, P^2)-L(f^-, P^2) \\ &< \varepsilon/2+\varepsilon/2 = \varepsilon \end{aligned}$$

y, de nuevo, según el teorema 2.14, f es integrable sobre $[\mathbf{a}, \mathbf{b}]$.

Antes de probar 3.5 en el caso general, lo probaremos para un caso especial.

3.12 Lema. *Si* $f(\mathbf{x}) \geqslant 0$ *para todo* $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$ *y* f *es integrable sobre* $[\mathbf{a}, \mathbf{b}]$, *entonces* f^2 *es integrable sobre* $[\mathbf{a}, \mathbf{b}]$.

PRUEBA. Sea P una partición de $[\mathbf{a}, \mathbf{b}]$. Para todo subintervalo $\mathscr{R}_i$,

$$M_i(f^2) = \sup\{f^2(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\} = (\sup\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\})^2 = M_i^2(f)$$

y

$$m_i(f^2) = \inf\{f^2(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\} = (\inf\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\})^2 = m_i^2(f).$$

Por tanto,

$$\begin{aligned} U(f^2, P)-L(f^2, P) &= \sum_{i=1}^{k} [M_i(f^2)-m_i(f^2)]\, A(\mathscr{R}_i) \\ &= \sum_{i=1}^{k} [M_i(f)+m_i(f)]\,[M_i(f)-m_i(f)]\, A(\mathscr{R}_i) \end{aligned}$$

$$\leqslant 2M \sum_{i=1}^{k} [M_i(f) - m_i(f)] A(\mathscr{R}_i)$$

$$= 2M[U(f, P) - L(f, P)].$$

Como f es integrable sobre $[\mathbf{a}, \mathbf{b}]$, según el teorema 2.14 ha de haber para cada $\varepsilon > 0$ una partición P de $[\mathbf{a}, \mathbf{b}]$ tal que el segundo miembro de la anterior desigualdad es menor que ε. Luego, de nuevo según el teorema 2.14 f^2 es integrable sobre $[\mathbf{a}, \mathbf{b}]$.

PRUEBA DE 3.5. Tenemos

$$f^2 = (f^+ - f^-)^2 = (f^+)^2 - 2f^+f^- + (f^-)^2 = (f^+)^2 + (f^-)^2$$

puesto que al menos uno de los dos, $f^+(\mathbf{x})$ o $f^-(\mathbf{x})$, es cero para todo $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$. Por 3.4, f^+ y f^- son integrables sobre $[\mathbf{a}, \mathbf{b}]$, luego, según el lema 3.12, $(f^+)^2$ y $(f^-)^2$ son integrables sobre $[\mathbf{a}, \mathbf{b}]$. La integrabilidad de f^2 se sigue entonces de 3.3.

PRUEBA DE 3.6. Según 3.3 y 3.2, $f+g$ y $f-g$ son integrables sobre $[\mathbf{a}, \mathbf{b}]$. Luego, por 3.5, $(f+g)^2$ y $(f-g)^2$ son integrables sobre $[\mathbf{a}, \mathbf{b}]$. La integrabilidad de fg se sigue entonces de 3.3 y 3.2 observando que

$$fg = \tfrac{1}{4}[f+g]^2 - \tfrac{1}{4}[f-g]^2.$$

PRUEBA DE 3.7. Es claro que $f(\mathbf{x}) \leqslant g(\mathbf{x})$ para todo $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$ implica $\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant \underline{\int_{\mathbf{a}}^{\mathbf{b}}} g$ y $\overline{\int_{\mathbf{a}}^{\mathbf{b}}} f \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} g$. Como f y g se han supuesto integrables, de cualquiera de las anteriores desigualdades se sigue que $\int_{\mathbf{a}}^{\mathbf{b}} f \leqslant \int_{\mathbf{a}}^{\mathbf{b}} g$.

PRUEBA DE 3.8. Como para todo $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$, $f^+(\mathbf{x}) \geqslant 0$ y $f^-(\mathbf{x}) \geqslant 0$, tenemos

$$f(\mathbf{x}) = f^+(\mathbf{x}) - f^-(\mathbf{x}) \leqslant f^+(\mathbf{x}) + f^-(\mathbf{x}) = |f(\mathbf{x})|$$

y

$$-f(\mathbf{x}) = -f^+(\mathbf{x}) + f^-(\mathbf{x}) \leqslant f^+(\mathbf{x}) + f^-(\mathbf{x}) = |f(\mathbf{x})|.$$

Por 3.4, f^+ y f^- son integrables sobre $[\mathbf{a}, \mathbf{b}]$, y por 3.3 $|f|$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$. Por 3.7 tenemos

$$\int_{\mathbf{a}}^{\mathbf{b}} f \leqslant \int_{\mathbf{a}}^{\mathbf{b}} |f| \quad \text{e} \quad -\int_{\mathbf{a}}^{\mathbf{b}} f \leqslant \int_{\mathbf{a}}^{\mathbf{b}} |f|.$$

Combinando estas dos desigualdades tenemos

$$\left| \int_{\mathbf{a}}^{\mathbf{b}} f \right| \leqslant \int_{\mathbf{a}}^{\mathbf{b}} |f|.$$

4. INTEGRALES SOBRE CONJUNTOS ACOTADOS EN R^2

En esta sección extendemos la definición de integral de f a conjuntos acotados más generales $\mathscr{E}$ en R^2.

4.1 Definición. *Si f es una función definida y acotada sobre un conjunto acotado $\mathscr{E} \subset R^2$, definimos la función $f_{\mathscr{E}}$ por la regla*

$$f_{\mathscr{E}}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \textit{para} \quad \mathbf{x} \in \mathscr{E} \\ 0 & \textit{para} \quad \mathbf{x} \in \mathscr{C}\mathscr{E} \end{cases}$$

donde $\mathscr{C}\mathscr{E}$, el complemento de $\mathscr{E}$, es el conjunto de todos los puntos de R^2 que no están en $\mathscr{E}$.

Nota. Recuérdese que hemos usado anteriormente la notación $f_{\mathscr{E}}$ para denotar la restricción de una función f a un conjunto $\mathscr{E} \subset \mathscr{D}_f$, es decir, $f_{\mathscr{E}}(\mathbf{x}) = f(\mathbf{x})$ para $\mathbf{x} \in \mathscr{E}$. En este capítulo usamos la misma notación para denotar la función que es igual a f sobre $\mathscr{E}$ y cero en cualquier otro punto.

4.2 Definición. *Si $\mathscr{E}$ es un conjunto acotado en R^2, si $[\mathbf{a}, \mathbf{b}]$ es un intervalo en R^2 tal que $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$, y si $\int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}}$ existe, entonces definimos*

$$\int_{\mathscr{E}} f = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}}.$$

El valor de $\int_{\mathscr{E}} f$ no depende de la elección del intervalo $[\mathbf{a}, \mathbf{b}]$. Si así no fuera esta integral no estaría definida. Para mostrar esto necesitamos primero probar que la intersección de dos intervalos es o un intervalo o el conjunto vacío. Sea

$$[\mathbf{a}^1, \mathbf{b}^1] = \{(x, y) \mid a_1{}^1 \leqslant x \leqslant b_1{}^1, a_2{}^1 \leqslant y \leqslant b_2{}^1\}$$

y

$$[\mathbf{a}^2, \mathbf{b}^2] = \{(x, y) \mid a_1{}^2 \leqslant x \leqslant b_1{}^2, a_2{}^2 \leqslant y \leqslant b_2{}^2\}.$$

Entonces

$$[\mathbf{a}^1, \mathbf{b}^2] \cap [\mathbf{a}^2, \mathbf{b}^2] = [\mathbf{a}^3, \mathbf{b}^3]$$

donde los componentes de $\mathbf{a}^3$ y $\mathbf{b}^3$ están definidos por

$$a_i{}^3 = \text{máx}\,\{a_i{}^1, a_i{}^2\} \quad \text{y} \quad b_i{}^3 = \text{mín}\,\{b_i{}^1, b_i{}^2\} \quad (i = 1, 2).$$

Si $b_1{}^3 < a_1{}^3$ o $b_2{}^3 < a_2{}^3$, entonces $[\mathbf{a}^3, \mathbf{b}^3] = \phi$.

Sean $[\mathbf{a}^1, \mathbf{b}^1]$ y $[\mathbf{a}^2, \mathbf{b}^2]$ dos intervalos que contienen $\mathscr{E}$. Entonces $\mathscr{E} \subset [\mathbf{a}^1, \mathbf{b}^1] \cap [\mathbf{a}^2, \mathbf{b}^2] = [\mathbf{a}^3, \mathbf{b}^3]$. Esta relación está ilustrada en la figura 6. Como $f_{\mathscr{E}}(\mathbf{x}) = 0$ para aquellos puntos de $[\mathbf{a}^1, \mathbf{b}^1]$ que no están en $[\mathbf{a}^3, \mathbf{b}^3]$

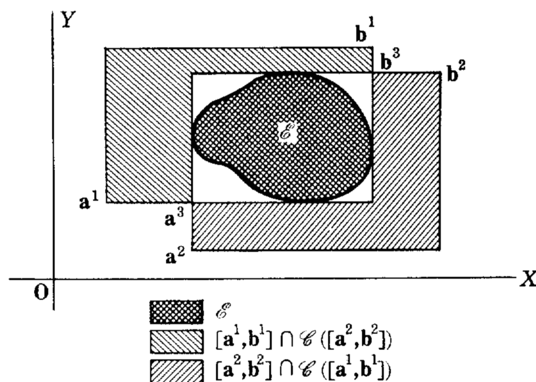


FIGURA 6

y para aquellos puntos de $[\mathbf{a}^2, \mathbf{b}^2]$ que no están en $[\mathbf{a}^3, \mathbf{b}^3]$,

$$\int_{\mathbf{a}^1}^{\mathbf{b}^1} f_{\mathscr{E}} = \int_{\mathbf{a}^3}^{\mathbf{b}^3} f_{\mathscr{E}} = \int_{\mathbf{a}^2}^{\mathbf{b}^2} f_{\mathscr{E}}.$$

4.3 Ejemplo. Sea $\mathscr{E} = \{\mathbf{x} \mid x \in [0, 1], y \in [0, 1], \text{ y } x, y \text{ racionales}\}$. Demuéstrese que $\int_{\mathscr{E}} 1$ no existe.

SOLUCIÓN. Tomemos $\mathbf{a} = (0, 0)$ y $\mathbf{b} = (1, 1)$. Entonces $\mathscr{E}$ es el conjunto de todos los puntos en $[\mathbf{a}, \mathbf{b}]$ con ambas coordenadas números racionales. La función $1_{\mathscr{E}}$ está definida por

$$1_{\mathscr{E}}(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \mathscr{E} \\ 0 & \mathbf{x} \in \mathscr{C}\mathscr{E}. \end{cases}$$

Sea P una partición de $[\mathbf{a}, \mathbf{b}]$. Como todo subintervalo $\mathscr{R}_i$ en la partición P contiene puntos con ambas coordenadas racionales al igual que puntos con, al menos, una de las coordenadas irracional,

$$m_i = \text{ínf } \{1_{\mathscr{E}}(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\} = 0$$

y

$$M_i = \sup \{1_{\mathscr{E}}(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\} = 1.$$

De donde $\overline{\int_{\mathbf{a}}^{\mathbf{b}}} 1_{\mathscr{E}} = 1$ e $\underline{\int_{\mathbf{a}}^{\mathbf{b}}} 1_{\mathscr{E}} = 0$, Por tanto, $\int_{\mathscr{E}} 1$ no existe.

4.4 Ejemplo. Sea f la función con regla de correspondencia

$$f(x, y) = \sqrt{1-(x^2+y^2)}$$

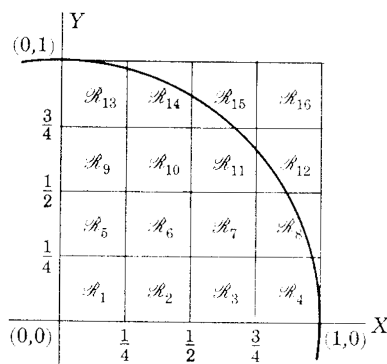


FIGURA 7

y dominio la región limitada por la circunferencia unitaria

$$\mathscr{E} = \{(x, y) : x^2 + y^2 \leqslant 1\}.$$

Suponiendo que f sea integrable sobre $\mathscr{E}$ (más adelante veremos que tal cosa es cierta), demuéstrese que

$$1 < \int_{\mathscr{E}} f < 3.$$

Solución. Como tanto $\mathscr{E}$ como f son, ambos, simétricos con respecto a los ejes X y Y, es suficiente estimar la integral sobre la parte de $\mathscr{E}$ situada sobre el primer cuadrante y multiplicar este resultado por 4. El rectángulo más pequeño que contiene esta parte de $\mathscr{E}$ en el primer cuadrante, es el cuadrado $\mathscr{S} = [(0, 0), (1, 1)]$ (figura 7).

Tomemos la partición P del cuadrado $\mathscr{S} = [(0, 0), (1, 1)]$ donde

$$P_1 = P_2 = \{0, \tfrac{1}{4}, \tfrac{1}{2}, \tfrac{3}{4}, 1\}$$

y numeremos las $4 \times 4 = 16$ subregiones $\mathscr{R}_i$ como se muestra en la figura 7. Los valores de m_i y M_i se exhiben en la siguiente tabla.

i	m_i	M_i	i	m_i	M_i
1	0.935	1.000	9	0.613	0.867
2	0.830	0.969	10	0.434	0.830
3	0.613	0.867	11	0.000	0.707
4	0.000	0.662	12	0.000	0.434
5	0.830	0.969	13	0.000	0.662
6	0.707	0.935	14	0.000	0.613
7	0.434	0.830	15	0.000	0.434
8	0.000	0.613	16	0.000	0.000

Tenemos

$$L(f, P) = 4 \sum_{i=1}^{16} m_i(f) A(\mathscr{R}_i) = \tfrac{4}{16} \sum_{i=1}^{16} m_i(f) = \tfrac{1}{4}(5.396) = 1.349$$

y

$$U(f, P) = 4 \sum_{i=1}^{16} M_i(f) A(\mathscr{R}_i) = \tfrac{1}{4} \sum_{i=1}^{16} M_i(f) = \tfrac{1}{4}(11.392) = 2.848.$$

Como $L(f, P) \leqslant 4 \int_{(0,0)}^{(1,1)} f_{\mathscr{E}} = \int_{\mathscr{E}} f \leqslant U(f, P)$, hemos establecido la desigualdad.

Podríamos esperar que $\frac{1}{2}[L(f, P)+U(f, P)] = 2.098$ sería una buena aproximación a $\int_{\mathscr{E}} f$. En realidad, como veremos más tarde [problema (*e*), pág. 357]

$$\int_{\mathscr{E}} f = \tfrac{2}{3}\pi \approx 2.094.$$

El cálculo del ejemplo 4.4 es realmente largo. Sin embargo, con las modernas calculadoras, conseguir una gran exactitud en el cálculo de tales integrales es una tarea sencilla.

4.5 Definición. *Si $\mathscr{E}$ es un conjunto acotado en* R^2 *y* $[\mathbf{a}, \mathbf{b}]$ *es un intervalo en* R^2 *tal que* $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$, *entonces*[1]

$$\underline{A}(\mathscr{E}) = \underline{\int_{\mathbf{a}}^{\mathbf{b}}} 1_{\mathscr{E}_i} \quad y \quad \bar{A}(\mathscr{E}) = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} 1_{\bar{\mathscr{E}}}$$

se llaman, respectivamente, ***área interior*** *y* ***área exterior*** *de $\mathscr{E}$. Si $\underline{A}(\mathscr{E}) = \bar{A}(\mathscr{E})$, entonces el valor común, denotado por $A(\mathscr{E})$, se llama* ***área*** *de $\mathscr{E}$.*

Puede probarse fácilmente que $\underline{A}(\mathscr{E})$ y $\bar{A}(\mathscr{E})$ no dependen de la elección del intervalo $[\mathbf{a}, \mathbf{b}]$ que contiene a $\mathscr{E}$.

Como demostraremos a continuación en el teorema 4.6, si $\mathscr{E}$ tiene área, entonces

$$A(\mathscr{E}) = \int_{\mathbf{a}}^{\mathbf{b}} 1_{\mathscr{E}}.$$

La función $1_{\mathscr{E}}$ se llama *función característica* del conjunto $\mathscr{E}$.

Cuando $\mathscr{E}$ es un intervalo, la definición 4.5 concuerda con la previamente dada para el área de un intervalo, pág. 313.

[1] Como es común, $\mathscr{E}_i$ denota el interior de $\mathscr{E}$, y $\bar{\mathscr{E}}$ denota la cerradura de $\mathscr{E}$. Véanse las páginas 164 y 166.

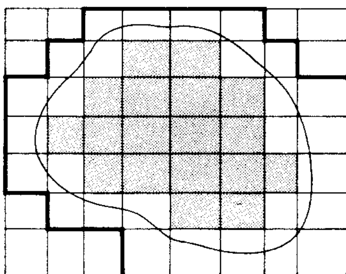


FIGURA 8

El área interior de $\mathscr{E}$, como integral inferior que es, es el supremo de sumas inferiores. Si P es una partición de $[\mathbf{a}, \mathbf{b}]$ y $\mathscr{R}_j$ denota el j-ésimo intervalo de P, entonces $m_j(1_{\mathscr{E}_i}) = 0$ para cada subintervalo $\mathscr{R}_j$ que contiene puntos de $\mathscr{E}_b \cup \mathscr{E}_e$ y, por tanto,

$$L(1_{\mathscr{E}_i}, P) = \sum_{j=1}^{k} m_j(1_{\mathscr{E}_i}) A(\mathscr{R}_j)$$

es la suma de las áreas de aquellos subintervalos de P que son subconjuntos de $\mathscr{E}_i$. Por otra parte, el área exterior de $\mathscr{E}$ es una integral superior y como tal es el ínfimo de sumas superiores. Ahora bien, $M_j(1_{\bar{\mathscr{E}}}) = 0$ solamente para aquellos subintervalos de P que no contienen ningún punto de $\mathscr{E}$, es decir, $M_j(1_{\bar{\mathscr{E}}}) = 0$ para aquellos subintervalos que contienen solamente puntos de $\mathscr{E}_e$, luego

$$U(1_{\bar{\mathscr{E}}}, P) = \sum_{j=1}^{k} M_j(1_{\bar{\mathscr{E}}}) A(\mathscr{R}_j)$$

es la suma de las áreas de aquellos subintervalos de P que contienen puntos de $\bar{\mathscr{E}}$. Vemos pues, que la suma inferior se aproxima al área de $\mathscr{E}$ por el área de un conjunto inscrito de rectángulos y las sumas superiores se aproximan al área de $\mathscr{E}$ por el área de un conjunto circunscrito de rectángulos. En la figura 8 las sumas superior e inferior están ilustradas para un conjunto $\mathscr{E}$ en R^2 donde $\mathscr{E}$ es el conjunto de todos los puntos en el interior y sobre la curva cerrada que allí se muestra. Los subintervalos que aparecen a la suma inferior están sombreados y los que se encuentran en el interior de la poligonal gruesa son los que pertenecen a la suma superior.

4.6 Teorema. *Si un conjunto acotado $\mathscr{E}$ en R^2 tiene área y $[\mathbf{a}, \mathbf{b}]$ es un intervalo tal que $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$, entonces*

$$A(\mathscr{E}) = \int_{\mathbf{a}}^{\mathbf{b}} 1_{\mathscr{E}}.$$

PRUEBA. Para todo $x \in [a, b]$, $0 \leqslant 1_{\mathscr{E}_1}(x) \leqslant 1_{\mathscr{E}}(x) \leqslant 1_{\bar{\mathscr{E}}}(x)$. Luego,

4.7 $$0 \leqslant \underline{A}(\mathscr{E}) = \underline{\int_a^b} 1_{\mathscr{E}_i} \leqslant \underline{\int_a^b} 1_{\mathscr{E}} \leqslant \overline{\int_a^b} 1_{\mathscr{E}} \leqslant \overline{\int_a^b} 1_{\bar{\mathscr{E}}} = \bar{A}(\mathscr{E}).$$

Como $\mathscr{E}$ tiene área, $\underline{A}(\mathscr{E}) = \bar{A}(\mathscr{E})$, y por tanto

$$A(\mathscr{E}) = \int_a^b 1_{\mathscr{E}}.$$

Demostraremos que el área tiene las siguientes propiedades fundamentales: si $\mathscr{E}$ y $\mathscr{F}$ son conjuntos en R^2 que tienen área, entonces

$$A(\mathscr{E}) \geqslant 0;$$

$$\text{si } \mathscr{E} \subset \mathscr{F}, \text{ entonces } A(\mathscr{E}) \leqslant A(\mathscr{F});$$

$$\text{si } A(\mathscr{E} \cap \mathscr{F}) = 0, \text{ entonces } A(\mathscr{E} \cup \mathscr{F}) = A(\mathscr{E}) + A(\mathscr{F}).$$

4.8 Teorema. *Si un conjunto acotado $\mathscr{E}$ en R^2 tiene área, entonces $0 \leqslant A(\mathscr{E})$.*

PRUEBA. El teorema se sigue inmediatamente de la desigualdad 4.7.

4.9 Lema. *Si $\mathscr{E}$ y $\mathscr{F}$ son conjuntos acotados en R^2 y $\mathscr{E} \subset \mathscr{F}$, entonces*

$$\underline{A}(\mathscr{E}) \leqslant \underline{A}(\mathscr{F}) \quad y \quad \bar{A}(\mathscr{E}) \leqslant \bar{A}(\mathscr{F}).$$

PRUEBA. Sea $[a, b]$ un intervalo en R^2 tal que $\mathscr{F} \subset [a, b]$. Entonces, como $\mathscr{E} \subset \mathscr{F}$, $1_{\mathscr{E}_i}(x) \leqslant 1_{\mathscr{F}_i}(x)$ y $1_{\bar{\mathscr{E}}}(x) \leqslant 1_{\bar{\mathscr{F}}}(x)$ para todo $x \in [a, b]$. De donde

$$\underline{\int_a^b} 1_{\mathscr{E}_i} \leqslant \underline{\int_a^b} 1_{\mathscr{F}_i} \quad \text{e} \quad \overline{\int_a^b} 1_{\bar{\mathscr{E}}} \leqslant \overline{\int_a^b} 1_{\bar{\mathscr{F}}}$$

o bien

$$\underline{A}(\mathscr{E}) \leqslant \underline{A}(\mathscr{F}) \quad \text{y} \quad \bar{A}(\mathscr{E}) \leqslant \bar{A}(\mathscr{F}).$$

4.10 Teorema. *Si dos conjuntos acotados $\mathscr{E}$ y $\mathscr{F}$ en R^2 tienen área y $\mathscr{E} \subset \mathscr{F}$, entonces*

$$A(\mathscr{E}) \leqslant A(\mathscr{F}).$$

PRUEBA. La desigualdad del teorema se sigue inmediatamente de cualquiera de las desigualdades del lema 4.9.

4.11 Lema. *Si $\mathscr{E}$ y $\mathscr{F}$ son conjuntos acotados en R^2, entonces*

$$\underline{A}(\mathscr{E} \cap \mathscr{F}) + \underline{A}(\mathscr{E} \cup \mathscr{F}) \geqslant \underline{A}(\mathscr{E}) + \underline{A}(\mathscr{F})$$

y

$$\bar{A}(\mathscr{E} \cap \mathscr{F}) + \bar{A}(\mathscr{E} \cup \mathscr{F}) \leqslant \bar{A}(\mathscr{E}) + \bar{A}(\mathscr{F}).$$

PRUEBA. Sea $[\mathbf{a}, \mathbf{b}]$ un intervalo en R^2 tal que $\mathscr{E} \cup \mathscr{F} \subset [\mathbf{a}, \mathbf{b}]$ y sea P una partición de $[\mathbf{a}, \mathbf{b}]$. Para cualquiera de los subintervalos $\mathscr{R}_j$ tenemos

$$m_j(1_{\mathscr{E}_i})+m_j(1_{\mathscr{F}_i}) \leqslant m_j(1_{\mathscr{E}_i \cap \mathscr{F}_i})+m_j(1_{\mathscr{E}_i \cup \mathscr{F}_i})$$

$$\leqslant m_j(1_{(\mathscr{E} \cap \mathscr{F})_i})+m_j(1_{(\mathscr{E} \cup \mathscr{F})_i})$$

como $\mathscr{E}_i \cap \mathscr{F}_i = (\mathscr{E} \cap \mathscr{F})_i$ y $\mathscr{E}_i \cup \mathscr{F}_i \subset (\mathscr{E} \cup \mathscr{F})_i$. Así pues, si P^1 y P^2 son particiones cualesquiera de $[\mathbf{a}, \mathbf{b}]$ y P es un refinamiento común de P^1 y P^2, entonces

$$L(1_{\mathscr{E}_i}, P^1)+L(1_{\mathscr{F}_i}, P^2) \leqslant L(1_{\mathscr{E}_i}, P)+L(1_{\mathscr{F}_i}, P)$$

$$\leqslant L(1_{(\mathscr{E} \cap \mathscr{F})_i}, P)+L(1_{(\mathscr{E} \cup \mathscr{F})_i}, P)$$

$$\leqslant \underline{A}(\mathscr{E} \cap \mathscr{F})+\underline{A}(\mathscr{E} \cup \mathscr{F}).$$

De donde resulta que para todo $P^1 \in \mathscr{P}$, $\underline{A}(\mathscr{E} \cap \mathscr{F})+\underline{A}(\mathscr{E} \cup \mathscr{F})-L(1_{\mathscr{E}_i}, P^1)$ es una cota superior de $\{L(1_{\mathscr{F}_i}, P^2) \mid P^2 \in \mathscr{P}\}$ y como $\underline{A}(\mathscr{F})$ es el supremo de este conjunto, tenemos

$$\underline{A}(\mathscr{F}) \leqslant \underline{A}(\mathscr{E} \cap \mathscr{F})+\underline{A}(\mathscr{E} \cup \mathscr{F})-L(1_{\mathscr{E}_i}, P^1)$$

o bien

$$L(1_{\mathscr{E}_i}, P^1) \leqslant \underline{A}(\mathscr{E} \cap \mathscr{F})+\underline{A}(\mathscr{E} \cup \mathscr{F})-\underline{A}(\mathscr{F}).$$

Luego $\underline{A}(\mathscr{E} \cap \mathscr{F})+\underline{A}(\mathscr{E} \cup \mathscr{F})-\underline{A}(\mathscr{F})$ es una cota superior de

$$\{L(1_{\mathscr{E}_i}, P^1) \mid P^1 \in \mathscr{P}\}$$

y como $\underline{A}(\mathscr{E})$ es el supremo de este conjunto, se tiene por ello la primera desigualdad del lema. Análogamente tenemos

$$M_j(1_{\overline{\mathscr{E}}})+M_j(1_{\overline{\mathscr{F}}}) \geqslant M_j(1_{\overline{\mathscr{E}} \cap \overline{\mathscr{F}}})+M_j(1_{\overline{\mathscr{E}} \cup \overline{\mathscr{F}}})$$

$$\geqslant M_j(1_{\overline{\mathscr{E} \cap \mathscr{F}}})+M_j(1_{\overline{\mathscr{E} \cup \mathscr{F}}})$$

ya que $\overline{\mathscr{E} \cap \mathscr{F}} \subset \overline{\mathscr{E}} \cap \overline{\mathscr{F}}$ y $\overline{\mathscr{E} \cup \mathscr{F}} = \overline{\mathscr{E}} \cup \overline{\mathscr{F}}$. De donde se obtiene la segunda desigualdad del lema.

4.12 Teorema. *Si dos conjuntos acotados $\mathscr{E}$ y $\mathscr{F}$ en R^2 tienen áreas, entonces $\mathscr{E} \cap \mathscr{F}$ y $\mathscr{E} \cup \mathscr{F}$ tienen área y*

$$A(\mathscr{E} \cap \mathscr{F})+A(\mathscr{E} \cup \mathscr{F}) = A(\mathscr{E})+A(\mathscr{F}).$$

PRUEBA. De acuerdo con el lema 4.11, como $\bar{A}(\mathscr{E}) = \underline{A}(\mathscr{E})$ y $\bar{A}(\mathscr{F}) = \underline{A}(\mathscr{F})$, tenemos

$$\underline{A}(\mathscr{E} \cap \mathscr{F})+\underline{A}(\mathscr{E} \cup \mathscr{F}) \geqslant A(\mathscr{E})+A(\mathscr{F}) \geqslant \bar{A}(\mathscr{E} \cap \mathscr{F})+\bar{A}(\mathscr{E} \cup \mathscr{F}).$$

Ninguno de los términos de la izquierda es mayor que el correspondiente

término de la derecha. De donde concluimos que $\underline{A}(\mathscr{E} \cap \mathscr{F}) = \bar{A}(\mathscr{E} \cap \mathscr{F})$ y $\underline{A}(\mathscr{E} \cup \mathscr{F}) = \bar{A}(\mathscr{E} \cup \mathscr{F})$ y

$$A(\mathscr{E} \cap \mathscr{F}) + A(\mathscr{E} \cup \mathscr{F}) = A(\mathscr{E}) + A(\mathscr{F}).$$

4.13 Corolario. *Si dos conjuntos acotados $\mathscr{E}$ y $\mathscr{F}$ en R^2 tienen área y $A(\mathscr{E} \cap \mathscr{F}) = 0$, entonces*

$$A(\mathscr{E} \cup \mathscr{F}) = A(\mathscr{E}) + A(\mathscr{F}).$$

Hemos probado las propiedades fundamentales del área: *si $\mathscr{E}$ y $\mathscr{F}$ son conjuntos en R^2 que tienen área, entonces*

(4.8′) $A(\mathscr{E}) \geqslant 0$;

(4.10′) *si* $\mathscr{E} \subset \mathscr{F}$, *entonces* $A(\mathscr{E}) \leqslant A(\mathscr{F})$;

(4.13′) *si* $A(\mathscr{E} \cap \mathscr{F}) = 0$, *entonces* $A(\mathscr{E} \cup \mathscr{F}) = A(\mathscr{E}) + A(\mathscr{F})$.

Es decir, el área es no negativa, el área de una parte no es mayor que el área del todo, y si un conjunto está dividido en dos partes que no se traslapan, la suma de las áreas de las partes es igual al área del conjunto.

Hacemos ahora la conexión entre área definida por integrales simples y área definida por integrales dobles. Si el área debe tener estas propiedades (4.8, 4.10 y 4.13), esto se demostró en el volumen I, páginas 543 a 545, que el área bajo la gráfica de una función f no negativa e integrable sobre un intervalo $[a, b]$ es $\int_a^b f$. Así pues, si una región tal tiene área en el sentido definido en esta sección —y probaremos en la sección 9 que la tiene— entonces las dos definiciones de área son las mismas para estas regiones. Así pues, la definición de área en términos de integrales dobles incluye la primera definición en términos de integrales simples. Sin embargo, la definición de área aquí dada, asigna área a conjuntos para los cuales la definición previa en términos de la integral simple no es aplicable. Es decir, la definición dada en esta sección amplía la clase de conjuntos de puntos en el plano a los que podemos asignar un área.

Si $\mathscr{A}$ y $\mathscr{B}$ son dos conjuntos, definimos el *conjunto diferencia* de $\mathscr{A}$ y $\mathscr{B}$, y denotamos por $\mathscr{A} - \mathscr{B}$, al conjunto de todos los elementos de $\mathscr{A}$ que no son elementos de $\mathscr{B}$, es decir,

$$\mathscr{A} - \mathscr{B} = \{x \mid x \in \mathscr{A} \text{ y } x \notin \mathscr{B}\}.$$

No requerimos que $\mathscr{B}$ sea un subconjunto de $\mathscr{A}$ para poder considerar la diferencia $\mathscr{A} - \mathscr{B}$.

Antes de probar un teorema concerniente al área de la diferencia de conjuntos observemos que $A([\mathbf{a}, \mathbf{b}]) - U(1_{\bar{\mathscr{E}}}, P)$ es la suma de las áreas

de los subintervalos de P que son subconjuntos de $(\mathscr{C}\mathscr{E})_i$ (figura 8). Por tanto,

$$A([\mathbf{a}, \mathbf{b}]) - U(1_{\bar{\mathscr{E}}}, P) = L(1_{(\mathscr{C}\mathscr{E})_i}, P).$$

Usamos este hecho en la prueba del siguiente lema.

4.14 Lema. *Si $\mathscr{E}$ es un conjunto acotado en* R^2 *que tiene área y* $[\mathbf{a}, \mathbf{b}]$ *es un intervalo en* R^2 *tal que* $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$, *entonces* $[\mathbf{a}, \mathbf{b}] - \mathscr{E} = \mathscr{C}_{[\mathbf{a},\mathbf{b}]}\mathscr{E}$ *tiene área y*

$$A(\mathscr{C}_{[\mathbf{a}.\mathbf{b}]}\mathscr{E}) = A([\mathbf{a}, \mathbf{b}]) - A(\mathscr{E}).$$

PRUEBA. Por brevedad denotaremos por $\mathscr{C}\mathscr{E}$ a $\mathscr{C}_{[\mathbf{a},\mathbf{b}]}\mathscr{E}$. Entonces

$$\begin{aligned} A([\mathbf{a}, \mathbf{b}]) - \bar{A}(\mathscr{E}) &= A([\mathbf{a}, \mathbf{b}]) - \inf\{U(1_{\bar{\mathscr{E}}}, P) \mid P \in \mathscr{P}\} \\ &= A([\mathbf{a}, \mathbf{b}]) + \sup\{-U(1_{\bar{\mathscr{E}}}, P) \mid P \in \mathscr{P}\} \\ &= \sup\{A([\mathbf{a}, \mathbf{b}]) - U(1_{\bar{\mathscr{E}}}, P) \mid P \in \mathscr{P}\} \\ &= \sup\{L(1_{(\mathscr{C}\mathscr{E})i}, P) \mid P \in \mathscr{P}\} = \underline{A}(\mathscr{C}\mathscr{E}). \end{aligned}$$

Análogamente

$$A([\mathbf{a}, \mathbf{b}]) - \underline{A}(\mathscr{E}) = \bar{A}(\mathscr{C}\mathscr{E}).$$

Como $\mathscr{E}$ tiene área, $\underline{A}(\mathscr{C}\mathscr{E}) = \bar{A}(\mathscr{C}\mathscr{E})$ y

$$A(\mathscr{C}\mathscr{E}) = A([\mathbf{a}, \mathbf{b}]) - A(\mathscr{E}).$$

4.15 Teorema. *Si dos conjuntos acotados $\mathscr{E}$ y $\mathscr{F}$ en* R^2 *tienen área, entonces $\mathscr{E} - \mathscr{F}$ tiene área y*

$$A(\mathscr{E} - \mathscr{F}) = A(\mathscr{E}) - A(\mathscr{E} \cap \mathscr{F}).$$

PRUEBA. Tomemos $[\mathbf{a}, \mathbf{b}]$ tal que $\mathscr{E} \cup \mathscr{F} \subset [\mathbf{a}, \mathbf{b}]$. Observemos primero que $\mathscr{E} - \mathscr{F} = \mathscr{E} \cap \mathscr{C}\mathscr{F}$ donde, por brevedad, hemos denotado a $\mathscr{C}_{[\mathrm{a},\mathrm{b}]}\mathscr{F}$ por $\mathscr{C}\mathscr{F}$. Por el lema 4.14, $\mathscr{C}\mathscr{F}$ tiene área. Como $\mathscr{E}$ y $\mathscr{C}\mathscr{F}$ tienen área, según el teorema 4.12, $\mathscr{E} - \mathscr{F} = \mathscr{E} \cap \mathscr{C}\mathscr{F}$ tiene área. Ahora bien, $(\mathscr{E} - \mathscr{F}) \cap (\mathscr{E} \cap \mathscr{F}) = \varnothing$ de modo que $A((\mathscr{E} - \mathscr{F}) \cap (\mathscr{E} \cap \mathscr{F})) = 0$ y de acuerdo con el corolario 4.13,

$$A(\mathscr{E}) = A((\mathscr{E} - \mathscr{F}) \cup (\mathscr{E} \cap \mathscr{F})) = A(\mathscr{E} - \mathscr{F}) + A(\mathscr{E} \cap \mathscr{F}).$$

4.16 Teorema. *Si un conjunto acotado $\mathscr{E}$ en* R^2 *tiene área, entonces $\mathscr{E}_i$ y $\bar{\mathscr{E}}$ tienen área y*

$$A(\mathscr{E}_i) = A(\mathscr{E}) = A(\bar{\mathscr{E}}).$$

PRUEBA. Como

$$(\mathscr{E}_i)_i = \mathscr{E}_i \subset (\bar{\mathscr{E}})_i \quad \text{y} \quad (\bar{\mathscr{E}}_i) \subset \bar{\mathscr{E}} = \bar{\bar{\mathscr{E}}},$$

tenemos

$$\underline{A}(\mathscr{E}_i) = \underline{A}(\mathscr{E}) \leqslant \underline{A}(\bar{\mathscr{E}}) \quad \text{y} \quad \bar{A}(\mathscr{E}_i) \leqslant \bar{A}(\mathscr{E}) = \bar{A}(\bar{\mathscr{E}}).$$

Así pues

$$\underline{A}(\mathscr{E}) = \underline{A}(\mathscr{E}_i) \leqslant \bar{A}(\mathscr{E}_i) \leqslant \bar{A}(\mathscr{E})$$

y

$$\underline{A}(\mathscr{E}) \leqslant \underline{A}(\bar{\mathscr{E}}) \leqslant \bar{A}(\bar{\mathscr{E}}) = \bar{A}(\mathscr{E}).$$

Como $\underline{A}(\mathscr{E}) = \bar{A}(\mathscr{E})$, tenemos $\underline{A}(\mathscr{E}_i) = \bar{A}(\mathscr{E}_i)$, $\underline{A}(\bar{\mathscr{E}}) = \bar{A}(\bar{\mathscr{E}})$, y

$$A(\mathscr{E}_i) = A(\mathscr{E}) = A(\bar{\mathscr{E}}).$$

4.17 Teorema. *Si un conjunto acotado $\mathscr{E}$ en* R^2 *tiene área, entonces la frontera de $\mathscr{E}$ tiene área y $A(\mathscr{E}_b) = 0$.*

PRUEBA. Como $\mathscr{E}$ tiene área, según el teorema 4.16, $\mathscr{E}$ y $\bar{\mathscr{E}}$ tienen área y $A(\mathscr{E}_i) = A(\mathscr{E}) = A(\bar{\mathscr{E}})$. De donde, de acuerdo con el teorema 4.15, tenemos

$$A(\mathscr{E}_b) = A(\bar{\mathscr{E}} - \mathscr{E}_i) = A(\bar{\mathscr{E}}) - A(\mathscr{E}_i) = 0.$$

Problemas

1. Supóngase —como es el caso— que cada una de las funciones que abajo aparecen es integrable en el conjunto que se indica. Demuéstrese que:

a) $0.2 \leqslant \int_{\mathscr{E}} f \leqslant 1.1$, si $f(x, y) = x + y$,

$$\mathscr{E} = \{(x, y) \mid 0 \leqslant x,\ 0 \leqslant y,\ x^2 + y^2 \leqslant 1\}$$

b) $0.008 \leqslant \int_{\mathscr{E}} f \leqslant 0.421$, si $f(x, y) = x^2 y^2$, $\mathscr{E} = \{(x, y) \mid x^2 + y^2 \leqslant 1\}$

2. Pruébese que:

a) $2 \leqslant A(\mathscr{E}) \leqslant \frac{15}{4}$, si $\mathscr{E} = \{(x, y) \mid x^2 + y^2 \leqslant 1\}$

b) $\frac{9}{2} \leqslant A(\mathscr{E}) \leqslant \frac{29}{4}$, si $\mathscr{E} = \left\{(x, y) \,\middle|\, \frac{x^2}{4} + y^2 \leqslant 1\right\}$.

5. EXISTENCIA DE FUNCIONES INTEGRABLES

El nùmero denotado por $\int_{\mathscr{E}} f$ y que se llama integral definida de f sobre $\mathscr{E}$ ha sido definido para funciones integrables. Pero, ¿qué son las funciones integrables? En esta sección demostraremos que si una función es acotada sobre [**a**, **b**] y si el conjunto de puntos sobre [**a**, **b**] donde es discontinua tiene área cero, entonces es integrable sobre [**a**, **b**]. Se probará en el capítulo 8 (pág. 478), que si una función f es continua sobre [**a**, **b**] entonces f

es acotada sobre $[\mathbf{a}, \mathbf{b}]$. De aquí podemos concluir que las funciones continuas sobre $[\mathbf{a}, \mathbf{b}]$ son integrables sobre $[\mathbf{a}, \mathbf{b}]$.

5.1 Teorema. *Sea f una función acotada sobre un intervalo* $[\mathbf{a}, \mathbf{b}]$ *tal que el conjunto $\mathscr{E}$ de puntos de discontinuidad de f sobre* $[\mathbf{a}, \mathbf{b}]$ *tenga área ce.o.* *Entonces la* $\int_{\mathbf{a}}^{\mathbf{b}} f$ *existe.*

PRUEBA. Como el conjunto $\mathscr{E}$ tiene área cero y por tanto área exterior cero, para cada $\varepsilon > 0$ hay una partición P' de $[\mathbf{a}, \mathbf{b}]$ tal que la unión $\mathscr{A}$ de todos los subintervalos de la partición P' que contienen puntos de $\mathscr{E}$ tiene área $A(\mathscr{A}) < \varepsilon$. La unión de los subintervalos no en $\mathscr{A}$ forma un conjunto cerrado y acotado $\mathscr{B}$ que no contiene punto alguno de $\mathscr{E}$. La función f es continua sobre $\mathscr{B}$ y como $\mathscr{B}$ es cerrado y acotado, de acuerdo con el teorema 7.6, pág. 488, f es uniformemente continua (definición 7.5, pág. 477) sobre $\mathscr{B}$. Así pues, hay una $\delta > 0$ tal que $\mathbf{x}^1$, $\mathbf{x}^2 \in \mathscr{B}$ y $|\mathbf{x}^1 - \mathbf{x}^2| < \sqrt{2\delta}$ implica $|f(\mathbf{x}^1) - f(\mathbf{x}^2)| < \varepsilon$.

Sea P un refinamiento de P' de norma $|P| < \delta$. Separando la diferencia entre las sumas superior e inferior correspondientes a P en las contribuciones de aquellos subintervalos en $\mathscr{A}$ y en las contribuciones de los subintervalos en $\mathscr{B}$, tenemos

$$U(f, P) - L(f, P)$$
$$= \sum_{\mathscr{R}_i \subset \mathscr{A}} [M_i(f) - m_i(f)]\, A(\mathscr{R}_i) + \sum_{\mathscr{R}_i \subset \mathscr{B}} [M_i(f) - m_i(f)]\, A(\mathscr{R}_i).$$

Como f es acotado sobre $[\mathbf{a}, \mathbf{b}]$ supongamos $m \leqslant f(\mathbf{x}) \leqslant M$ para todo $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$. Entonces

$$\sum_{\mathscr{R}_i \subset \mathscr{A}} [M_i(f) - m_i(f)]\, A(\mathscr{R}_i) \leqslant \sum_{\mathscr{R}_i \subset \mathscr{A}} [M - m]\, A(\mathscr{R}_i)$$
$$= [M - m]\, A(\mathscr{A}) < [M - m]\, \varepsilon.$$

Como $|P| < \delta$, $|f(\mathbf{x}^1) - f(\mathbf{x}^2)| < \varepsilon$ siempre que $\mathbf{x}^1$, $\mathbf{x}^2 \in \mathscr{R}_i \subset \mathscr{B}$. Esto implica $M_i(f) - m_i(f) < \varepsilon$ siempre que $\mathscr{R}_i \subset \mathscr{B}$ y por tanto

$$\sum_{\mathscr{R}_i \subset \mathscr{B}} [M_i(f) - m_i(f)]\, A(\mathscr{R}_i) < \sum_{\mathscr{R}_i \subset \mathscr{B}} \varepsilon A(\mathscr{R}_i) \leqslant \varepsilon A([\mathbf{a}, \mathbf{b}]).$$

Por tanto

$$U(f, P) - L(f, P)$$
$$= \sum_{\mathscr{R}_i \subset \mathscr{A}} [M_i(f) - m_i(f)]\, A(\mathscr{R}_i) + \sum_{\mathscr{R}_i \subset \mathscr{B}} [M_i(f) - m_i(f)]\, A(\mathscr{R}_i)$$
$$< \{M - m + A([\mathbf{a}, \mathbf{b}])\}\, \varepsilon$$

y según el teorema 2.14, pág. 320 —con ε reemplazada por $\{M-m+A([\mathbf{a},\mathbf{b}])\}\varepsilon$— se sigue que $\int_{\mathbf{a}}^{\mathbf{b}} f$ existe.

Enumeramos ahora diversos resultados que son corolarios del teorema 5.1. El primer corolario demuestra que el valor de la integral $\int_{\mathbf{a}}^{\mathbf{b}} f$ es independiente de la elección de los valores de f sobre un subconjunto de $[\mathbf{a}, \mathbf{b}]$ de área cero.

5.2 Corolario. *Sean f y g dos funciones acotadas sobre un intervalo* $[\mathbf{a}, \mathbf{b}]$ *y continuas sobre* $[\mathbf{a}, \mathbf{b}]-\mathscr{E}$ *donde $\mathscr{E}$ es un subconjunto de* $[\mathbf{a}, \mathbf{b}]$ *que tiene área cero. Si $f = g$ sobre* $[\mathbf{a}, \mathbf{b}]-\mathscr{E}$, *entonces*

$$\int_{\mathbf{a}}^{\mathbf{b}} f = \int_{\mathbf{a}}^{\mathbf{b}} g .$$

PRUEBA. Como f y g están acotadas sobre $[\mathbf{a}, \mathbf{b}]$, existen números m y M tales que $m \leqslant f(\mathbf{x}) \leqslant M$ y $m \leqslant g(\mathbf{x}) \leqslant M$ para toda $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$. Definamos las particiones P', P y el conjunto $\mathscr{A}$ en igual forma que en el teorema 5.1. Entonces,

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} (f-g) \geqslant L(f-g, P) = \sum_{\mathscr{R}_i \subset \mathscr{A}} m_i(f-g) A(\mathscr{R}_i) \geqslant (m-M)\varepsilon$$

e

$$\overline{\int_{\mathbf{a}}^{\mathbf{b}}} (f-g) \leqslant U(f-g, P) = \sum_{\mathscr{R}_i \subset \mathscr{A}} M_i(f-g) A(\mathscr{R}_i) \leqslant (M-m)\varepsilon .$$

Así pues

$$-(M-m)\varepsilon \leqslant \underline{\int_{\mathbf{a}}^{\mathbf{b}}} (f-g) \leqslant \overline{\int_{\mathbf{a}}^{\mathbf{b}}} (f-g) \leqslant (M-m)\varepsilon$$

y como $\varepsilon > 0$ es arbitraria, tenemos $\int_{\mathbf{a}}^{\mathbf{b}} (f-g) = 0$. Según el teorema 5.1, f y g son integrables sobre $[\mathbf{a}, \mathbf{b}]$. Por tanto

$$\int_{\mathbf{a}}^{\mathbf{b}} f - \int_{\mathbf{a}}^{\mathbf{b}} g = \int_{\mathbf{a}}^{\mathbf{b}} (f-g) = 0 .$$

El siguiente corolario es el recíproco del teorema 4.17, pág. 337.

5.3 Corolario. *Un conjunto acotado $\mathscr{E}$ en* R^2 *tiene área si la frontera de $\mathscr{E}$ tiene área cero.*

PRUEBA. Sea $[\mathbf{a}, \mathbf{b}]$ un intervalo que contiene a $\mathscr{E}$. Las funciones $1_{\mathscr{E}_i}$ y $1_{\bar{\mathscr{E}}}$ son continuas e iguales en todos los puntos de $[\mathbf{a}, \mathbf{b}]-\mathscr{E}_b$. Luego, según el corolario 5.2,

$$\underline{A}(\mathscr{E}) = \int_{\mathbf{a}}^{\mathbf{b}} 1_{\mathscr{E}_i} = \int_{\mathbf{a}}^{\mathbf{b}} 1_{\bar{\mathscr{E}}} = \bar{A}(\mathscr{E}).$$

5.4 Corolario. *Sea $\mathscr{E}$ un conjunto acotado en R^2 que tiene área. Si f es acotada sobre $\mathscr{E}$ y continua en el interior de $\mathscr{E}$, entonces $\int_{\mathscr{E}} f$ existe.*

PRUEBA. La función $f_{\mathscr{E}}$ es acotada y puede tener puntos de discontinuidad solamente en puntos de la frontera de $\mathscr{E}$. Sea $[\mathbf{a}, \mathbf{b}]$ un intervalo en R^2 tal que $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$. Entonces el conjunto de puntos de discontinuidad de $f_{\mathscr{E}}$ sobre $[\mathbf{a}, \mathbf{b}]$ está contenido en la frontera de $\mathscr{E}$ y como, según el teorema 4.17, esta frontera tiene área cero, de acuerdo con el teorema 5.1

$$\int_{\mathscr{E}} f = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}}$$

existe.

La pequeña generalización siguiente, del corolario 5.4, resulta a veces útil.

5.5 Corolario. *Sea $\mathscr{E}$ un conjunto acotado en R^2 que tiene área. Si f es acotada sobre $\mathscr{E}$ y continua en el interior de $\mathscr{E}$, excepto sobre un conjunto $\mathscr{F}$ de área cero, entonces $\int_{\mathscr{E}} f$ existe.*

PRUEBA. Los puntos de discontinuidad de $f_{\mathscr{E}}$ están contenidos en $\mathscr{E}_b \cup \mathscr{F}$. Como $\mathscr{F}$ y $\mathscr{E}$ tienen, ambos, área cero, su unión tiene área cero. Sea $[\mathbf{a}, \mathbf{b}] \subset \mathrm{R}^2$ tal que $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$. Entonces, según el teorema 5.1,

$$\int_{\mathscr{E}} f = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}}$$

existe.

5.6 Corolario. *Sea $\mathscr{E}$ un conjunto cerrado y acotado en R^2 que tiene área. Si f es continua sobre $\mathscr{E}$, entonces $\int_{\mathscr{E}} f$ existe.*

PRUEBA. Como $\mathscr{E}$ es cerrado y acotado, f continua sobre $\mathscr{E}$ implica que f está acotada sobre $\mathscr{E}$ (teorema 7.7, pág. 488). Luego conforme al corolario 5.4, $\int_{\mathscr{E}} f$ existe.

Problemas

1. Pruébese que los siguientes conjuntos tienen área cero:

a) un número finito de puntos,
b) un segmento rectilíneo,
c) un número finito de segmentos rectilíneos.

2. Pruébese que un disco circular tiene área.

***3.** Supongamos que f es diferenciable sobre $[\mathbf{a}, \mathbf{b}]$ y que $|\mathbf{D}f(\mathbf{x})| \leqslant K$ para todo $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$. Aplíquese el teorema 6.10, pág. 194, y dedúzcase que: 1) para toda partición P, $U(f, P) - L(f, P) \leqslant \sqrt{2}\,|P|\,KA([\mathbf{a}, \mathbf{b}])$; 2) f es integrable sobre $[\mathbf{a}, \mathbf{b}]$; y 3) $\left|\int_{\mathbf{a}}^{\mathbf{b}} f - \frac{1}{2}(U(f, P) + L(f, P))\right| \leqslant \frac{\sqrt{2}}{2}|P|\,KA([\mathbf{a}, \mathbf{b}])$.

***4.** ¿Cuán pequeño tendría que hacerse $|P|$ para poder estar seguro de que el error en la aproximación de cada una de las siguientes integrales por sumas superiores e inferiores era menor que 0.0005 ?

a) $\displaystyle\iint_{(0,0)}^{(1,1)} (x+y)\,dx\,dy$ *b*) $\displaystyle\iint_{(0,0)}^{\left(\frac{\pi}{2},\frac{\pi}{2}\right)} (\operatorname{sen} x + \cos y)\,dx\,dy.$

6. PROPIEDADES BÁSICAS DE $\int_{\mathscr{E}} f$

Estableceremos ahora algunas propiedades básicas de la integral doble $\int_{\mathscr{E}} f$ donde $\mathscr{E}$ es un conjunto acotado en R^2. Estas propiedades son generalizaciones de las propiedades de $\int_{\mathbf{a}}^{\mathbf{b}} f$ dadas en la sección 3.

6.1 *Si la función f es integrable sobre un conjunto acotado $\mathscr{E}$ y c es una función constante de* R^2 *en* R, *entonces la función cf es integrable sobre $\mathscr{E}$ e* $\int_{\mathscr{E}} cf = c\int_{\mathscr{E}} f$.

6.2 *Si c es una función constante de* R^2 *en* R *y $\mathscr{E}$ es un conjunto acotado que tiene área, entonces c es integrable sobre $\mathscr{E}$ e* $\int_{\mathscr{E}} c = cA(\mathscr{E})$.

6.3 *Si las funciones f y g son integrables sobre un conjunto acotado $\mathscr{E}$, entonces la función $f+g$ es integrable sobre $\mathscr{E}$ e* $\int_{\mathscr{E}} (f+g) = \int_{\mathscr{E}} f + \int_{\mathscr{E}} g$.

6.4 *La función f es integrable sobre un conjunto acotado $\mathscr{E}$ si y sólo si las funciones f^+ y f^- definidas por las reglas*

$$f^+(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & si\ f(\mathbf{x}) \geqslant 0 \\ 0 & si\ f(\mathbf{x}) < 0 \end{cases} \quad y \quad f^-(\mathbf{x}) = \begin{cases} 0 & si\ f(\mathbf{x}) \geqslant 0 \\ -f(\mathbf{x}) & si\ f(\mathbf{x}) < 0 \end{cases}$$

son, ambas, integrables sobre $\mathscr{E}$.

6.5 *Si la función f es integrable sobre un conjunto acotado $\mathscr{E}$, entonces f^2 es integrable sobre $\mathscr{E}$.*

6.6 *Si las funciones f y g son integrables sobre un conjunto acotado $\mathscr{E}$, entonces el producto fg es integrable sobre $\mathscr{E}$.*

6.7 *Si las funciones f y g son integrables sobre un conjunto acotado $\mathscr{E}$ y $f(\mathbf{x}) \leqslant g(\mathbf{x})$ para toda $\mathbf{x} \in \mathscr{E}$, entonces $\displaystyle\int_{\mathscr{E}} f \leqslant \int_{\mathscr{E}} g$.*

6.8 *Si la función f es integrable sobre un conjunto acotado $\mathscr{E}$, entonces la función $|f|$ es integrable sobre $\mathscr{E}$ e $\displaystyle\left|\int_{\mathscr{E}} f\right| \leqslant \int_{\mathscr{E}} |f|$.*

PRUEBA. Estas propiedades se siguen directamente de las propiedades 3.1 a 3.8 de $\displaystyle\int_{\mathbf{a}}^{\mathbf{b}} f$, pág. 323, y de la definición 4.2 de $\displaystyle\int_{\mathscr{E}} f$, pág. 328. Sea $[\mathbf{a}, \mathbf{b}]$ un intervalo en R^2 tal que $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$.

(6.1) De acuerdo con la propiedad 3.1,

$$\int_{\mathscr{E}} cf = \int_{\mathbf{a}}^{\mathbf{b}} (cf)_{\mathscr{E}} = \int_{\mathbf{a}}^{\mathbf{b}} cf_{\mathscr{E}} = c\int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}} = c\int_{\mathscr{E}} f.$$

(6.2) Reemplazando f por 1 en 6.1, tenemos

$$\int_{\mathscr{E}} c = c\int_{\mathscr{E}} 1 = cA(\mathscr{E}).$$

(6.3) Según la propiedad 3.3,

$$\int_{\mathscr{E}} (f+g) = \int_{\mathbf{a}}^{\mathbf{b}} (f+g)_{\mathscr{E}} = \int_{\mathbf{a}}^{\mathbf{b}} (f_{\mathscr{E}}+g_{\mathscr{E}}) = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}} + \int_{\mathbf{a}}^{\mathbf{b}} g_{\mathscr{E}} = \int_{\mathscr{E}} f + \int_{\mathscr{E}} g.$$

(6.4) Como $f_{\mathscr{E}} = (f^+ - f^-)_{\mathscr{E}} = f_{\mathscr{E}}^+ - f_{\mathscr{E}}^-$, por la propiedad 3.4, $f_{\mathscr{E}}$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$ si y sólo si $f_{\mathscr{E}}^+$ y $f_{\mathscr{E}}^-$ son integrables sobre $[\mathbf{a}, \mathbf{b}]$ y

$$\int_{\mathscr{E}} f = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}} = \int_{\mathbf{a}}^{\mathbf{b}} (f_{\mathscr{E}}^+ - f_{\mathscr{E}}^-) = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}}^+ - \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}}^- = \int_{\mathscr{E}} f^+ - \int_{\mathscr{E}} f^-.$$

(6.5) Según la definición 4.2, f es integrable sobre $\mathscr{E}$ si $f_{\mathscr{E}}$ es integrable

sobre [**a**, **b**]. Por la propiedad 3.5, $f_{\mathscr{E}}$ integrable sobre [**a**, **b**] implica $f_{\mathscr{E}}^2$ integrable sobre [**a**, **b**], y de aquí, de nuevo, por la definición 4.2, resulta que f^2 es integrable sobre $\mathscr{E}$.

(6.6) Según la definición 4.2, f y g son integrables sobre $\mathscr{E}$ si $f_{\mathscr{E}}$ y $g_{\mathscr{E}}$ son integrables sobre [**a**, **b**]. Según la propiedad 3.6 esto implica que $f_{\mathscr{E}} g_{\mathscr{E}} = (fg)_{\mathscr{E}}$ es integrable sobre [**a**, **b**] y, de nuevo, por la definición 4.2 esto, a su vez, implica que fg es integrable sobre $\mathscr{E}$.

(6.7) $f(\mathbf{x}) \leqslant g(\mathbf{x})$ para todo $\mathbf{x} \in \mathscr{E}$ implica $f_{\mathscr{E}}(\mathbf{x}) \leqslant g_{\mathscr{E}}(\mathbf{x})$ para toda $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$. De donde, de acuerdo con la propiedad 3.7, tenemos

$$\int_{\mathscr{E}} f = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}} \leqslant \int_{\mathbf{a}}^{\mathbf{b}} g_{\mathscr{E}} = \int_{\mathscr{E}} g \, .$$

(6.8) Según la definición 4.2, f es integrable sobre $\mathscr{E}$ si $f_{\mathscr{E}}$ es integrable sobre [**a**, **b**]. Pero $|f_{\mathscr{E}}| = |f|_{\mathscr{E}}$ y, por tanto, de nuevo, según 4.2, tenemos

$$\left| \int_{\mathscr{E}} f \right| = \left| \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}} \right| \leqslant \int_{\mathbf{a}}^{\mathbf{b}} |f_{\mathscr{E}}| = \int_{\mathbf{a}}^{\mathbf{b}} |f|_{\mathscr{E}} = \int_{\mathscr{E}} |f| \, .$$

6.9 Teorema. *Si f es integrable sobre un conjunto acotado $\mathscr{E}$ que tiene área, entonces*

$$\int_{\mathscr{E}} f = NA(\mathscr{E})$$

donde N es algún número entre m = ínf $\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{E}\}$ y M = sup $f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{E}\}$.

PRUEBA. Conforme a las propiedades 6.2 y 6.7

$$mA(\mathscr{E}) = \int_{\mathscr{E}} m \leqslant \int_{\mathscr{E}} f \leqslant \int_{\mathscr{E}} M = MA(\mathscr{E}).$$

De donde el teorema se sigue.

Nota. Si f es continua sobre $\mathscr{E}$, $f(\mathbf{x}) > m$ para todo $\mathbf{x} \in \mathscr{E}$, y $A(\mathscr{E}) > 0$, entonces se verifica la desigualdad estricta

$$mA(\mathscr{E}) < \int_{\mathscr{E}} f$$

y $N > m$. De modo análogo, si $f(\mathbf{x}) < M$ para toda $\mathbf{x} \in \mathscr{E}$ y $A(\mathscr{E}) > 0$, entonces $N < M$.

6.10 Corolario. *(Teorema del valor medio para integrales). Si f es continua*

y acotada sobre un conjunto $\mathscr{E}$ que tiene área y $\mathscr{E}_i$ es conexo, entonces existe un punto $\mathbf{x}_0 \in \mathscr{E}$ *tal que*

$$\int_{\mathscr{E}} f = f(\mathbf{x}_0) A(\mathscr{E}).$$

PRUEBA. Como $\mathscr{E}$ tiene área, $A(\mathscr{E}_b) = 0$ y $A(\mathscr{E}) = A(\mathscr{E}_i)$. Sea $[\mathbf{a}, \mathbf{b}] \subset \mathrm{R}^2$ un intervalo que contiene a $\mathscr{E}$. Por el corolario 5.2

$$\int_{\mathscr{E}} f = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}} = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}_i} = \int_{\mathscr{E}_i} f.$$

Según el teorema 6.9

$$\int_{\mathscr{E}} f = NA(\mathscr{E}_i)$$

donde N se encuentra entre $m = \text{ínf}\,\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{E}_i\}$ y $M = \sup\,\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{E}_i\}$. Entonces, existe un punto $a \in \mathscr{E}$ tal que $f(\mathbf{a}) = m$ o, en otro caso, $f(\mathbf{x}) > m$ para todo $\mathbf{x} \in \mathscr{E}_i$ y, según la observación anterior, existe un punto $\mathbf{a} \in \mathscr{E}_i$ tal que $m < f(\mathbf{a}) \leqslant N$. Análogamente, existe un punto $b \in \mathscr{E}_i$ tal que $N \leqslant f(\mathbf{b})$. Luego, de conformidad con el teorema del valor intermedio (pág. 187) existe un punto $\mathbf{x}_0 \in \mathscr{E}_i$ tal que $f(\mathbf{x}_0) = N$. De donde

$$\int_{\mathscr{E}} f = \int_{\mathscr{E}_i} f = NA(\mathscr{E}_i) = f(\mathbf{x}_0) A(\mathscr{E}).$$

Antes de proseguir estableciendo nuevas propiedades de $\int_{\mathscr{E}} f$, probamos el siguiente lema.

6.11 Lema. *Si f es acotada sobre un conjunto $\mathscr{E}$ y $A(\mathscr{E}) = 0$, entonces f es integrable sobre $\mathscr{E}$ e* $\int_{\mathscr{E}} f = 0$.

PRUEBA. De acuerdo al corolario 5.5 f es integrable sobre $\mathscr{E}$. Como f es acotada sobre $\mathscr{E}$, hay un número M tal que $|f(\mathbf{x}) \leqslant M$ para todo $\mathbf{x} \in \mathscr{E}$. Según las propiedades 6.8, 6.7 y 6.2

$$\left| \int_{\mathscr{E}} f \right| \leqslant \int_{\mathscr{E}} |f| \leqslant \int_{\mathscr{E}} M = MA(\mathscr{E}) = 0.$$

Para las integrales simples sabemos que si una función f es integrable sobre los intervalos $[a, b]$ y $[b, c]$ entonces f es integrable sobre $[a, c]$ e $\int_a^c f = \int_a^b f + \int_b^c f$. Probaremos ahora que las integrales dobles satisfacen una relación análoga.

6.12 Teorema. *Si f es integrable sobre los conjuntos acotados $\mathscr{E}_1$ y $\mathscr{E}_2$ y si*

al menos uno de los conjuntos $\mathscr{E}_1$, $\mathscr{E}_2$, o $\mathscr{E}_1 \cap \mathscr{E}_2$ tiene área, entonces f es integrable sobre la intersección $\mathscr{E}_1 \cap \mathscr{E}_2$ y la unión $\mathscr{E}_1 \cup \mathscr{E}_2$ e

$$\int_{\mathscr{E}_1 \cap \mathscr{E}_2} f + \int_{\mathscr{E}_1 \cup \mathscr{E}_2} f = \int_{\mathscr{E}_1} f + \int_{\mathscr{E}_2} f.$$

PRUEBA. Sea $[\mathbf{a}, \mathbf{b}]$ un intervalo en R^2 tal que $\mathscr{E}_1 \cup \mathscr{E}_2 \subset [\mathbf{a}, \mathbf{b}]$. Claramente, $f_{\mathscr{E}_1 \cap \mathscr{E}_2} = f_{\mathscr{E}_1} \cdot 1_{\mathscr{E}_1 \cap \mathscr{E}_2} = f_{\mathscr{E}_1} \cdot 1_{\mathscr{E}_2} = f_{\mathscr{E}_2} \cdot 1_{\mathscr{E}_1}$. Como f es integrable sobre $\mathscr{E}_1$ y $\mathscr{E}_2$, $f_{\mathscr{E}_1}$ y $f_{\mathscr{E}_2}$ son integrables sobre $[\mathbf{a}, \mathbf{b}]$. Por otra parte, al menos una de las funciones $1_{\mathscr{E}_1 \cap \mathscr{E}_2}$, $1_{\mathscr{E}_1}$, o $1_{\mathscr{E}_2}$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$. De donde, por la propiedad 3.6, $f_{\mathscr{E}_1 \cap \mathscr{E}_2}$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$; es decir, f es integrable sobre $\mathscr{E}_1 \cap \mathscr{E}_2$.

Probamos a continuación que

$$f_{\mathscr{E}_1 \cup \mathscr{E}_2} = f_{\mathscr{E}_1} + f_{\mathscr{E}_2} - f_{\mathscr{E}_1 \cap \mathscr{E}_2}.$$

Si $\mathbf{x} \in \mathscr{E}_1 \cup \mathscr{E}_2$, entonces $\mathbf{x} \in \mathscr{E}_1$ o $\mathbf{x} \in \mathscr{E}_2$. Pero $\mathbf{x} \in \mathscr{E}_1$ implica $f_{\mathscr{E}_1 \cup \mathscr{E}_2}(\mathbf{x}) = f(\mathbf{x}) = f_{\mathscr{E}_1}(\mathbf{x})$ y $f_{\mathscr{E}_2}(\mathbf{x}) = f_{\mathscr{E}_1 \cap \mathscr{E}_2}(\mathbf{x})$ mientras que $\mathbf{x} \in \mathscr{E}_2$ implica $f_{\mathscr{E}_1 \cup \mathscr{E}_2}(\mathbf{x}) = f(\mathbf{x}) = f_{\mathscr{E}_2}(\mathbf{x})$ y $f_{\mathscr{E}_1}(\mathbf{x}) = f_{\mathscr{E}_1 \cap \mathscr{E}_2}(\mathbf{x})$. Así pues, si $\mathbf{x} \in \mathscr{E}_1 \cup \mathscr{E}_2$, entonces $f_{\mathscr{E}_1 \cup \mathscr{E}_2}(\mathbf{x}) = f_{\mathscr{E}_1}(\mathbf{x}) + f_{\mathscr{E}_2}(\mathbf{x}) - f_{\mathscr{E}_1 \cap \mathscr{E}_2}(\mathbf{x})$. Si $\mathbf{x} \notin \mathscr{E}_1 \cup \mathscr{E}_2$, entonces $f_{\mathscr{E}_1}(\mathbf{x}) = f_{\mathscr{E}_2}(\mathbf{x}) = f_{\mathscr{E}_1 \cup \mathscr{E}_2}(\mathbf{x}) = f_{\mathscr{E}_1 \cap \mathscr{E}_2}(\mathbf{x}) = 0$ y de nuevo la fórmula se verifica.

Como $f_{\mathscr{E}_1}$, $f_{\mathscr{E}_2}$ y $f_{\mathscr{E}_1 \cap \mathscr{E}_2}$ son integrables sobre $[\mathbf{a}, \mathbf{b}]$ de acuerdo con las propiedades 3.3 y 3.1, tenemos

$$\int_{\mathscr{E}_1 \cup \mathscr{E}_2} f = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}_1 \cup \mathscr{E}_2} = \int_{\mathbf{a}}^{\mathbf{b}} (f_{\mathscr{E}_1} + f_{\mathscr{E}_2} - f_{\mathscr{E}_1 \cap \mathscr{E}_2})$$

$$= \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}_1} + \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}_2} - \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}_1 \cap \mathscr{E}_2} = \int_{\mathscr{E}_1} f + \int_{\mathscr{E}_2} f - \int_{\mathscr{E}_1 \cap \mathscr{E}_2} f.$$

6.13 Corolario. *Si f es integrable sobre los conjuntos acotados $\mathscr{E}_1$ y $\mathscr{E}_2$ y si $A(\mathscr{E}_1 \cap \mathscr{E}_2) = 0$, entonces f es integrable sobre $\mathscr{E}_1 \cup \mathscr{E}_2$ e*

$$\int_{\mathscr{E}_1 \cup \mathscr{E}_2} f = \int_{\mathscr{E}_1} f + \int_{\mathscr{E}_2} f.$$

PRUEBA. Según el teorema 6.12 f es integrable sobre $\mathscr{E}_1 \cap \mathscr{E}_2$ y, por tanto, acotada sobre $\mathscr{E}_1 \cap \mathscr{E}_2$. Luego, por el lema 6.11, $\int_{\mathscr{E}_1 \cap \mathscr{E}_2} f = 0$ y el corolario se sigue del teorema 6.12.

Usaremos para probar un recíproco parcial el teorema 6.12.

6.14 Corolario. *Si $\mathscr{E}_1$ y $\mathscr{E}_2$ son conjuntos que tienen área y la función f es integrable sobre $\mathscr{E}_1 \cup \mathscr{E}_2$, entonces f es integrable sobre $\mathscr{E}_1$, $\mathscr{E}_2$ y $\mathscr{E}_1 \cap \mathscr{E}_2$ e*

$$\int_{\mathscr{E}_1} f + \int_{\mathscr{E}_2} f - \int_{\mathscr{E}_1 \cap \mathscr{E}_2} f = \int_{\mathscr{E}_1 \cup \mathscr{E}_2} f.$$

PRUEBA. Sea [**a**, **b**] un intervalo en R^2 tal que $\mathscr{E}_1 \cup \mathscr{E}_2 \subset$ [**a**, **b**]. Tenemos

$$f_{\mathscr{E}_1} = f_{\mathscr{E}_1 \cup \mathscr{E}_2} \cdot 1_{\mathscr{E}_1} \quad \text{y} \quad f_{\mathscr{E}_2} = f_{\mathscr{E}_1 \cup \mathscr{E}_2} \cdot 1_{\mathscr{E}_2}$$

y, como $f_{\mathscr{E}_1 \cup \mathscr{E}_2}$, $1_{\mathscr{E}_1}$, y $1_{\mathscr{E}_2}$ son integrables sobre [**a**, **b**] por la propiedad 3.6 $f_{\mathscr{E}_1}$ y $f_{\mathscr{E}_2}$ son integrables sobre [**a**, **b**]; es decir, f es integrable sobre $\mathscr{E}_1$ y $\mathscr{E}_2$. Los conjuntos $\mathscr{E}_1$ y $\mathscr{E}_2$ tienen área, de forma que el teorema 6.12 es ahora aplicable con lo que se prueba el corolario.

7. INTEGRALES ITERADAS

Una integral de la forma

7.1
$$\int_a^b \int_{g(x)}^{h(x)} f(x,y)\,dy\,dx$$

se llama *integral iterada*. La integral 7.1 se interpreta como $\int_a^b F(x)\,dx$ donde para cada $x \in [a, b]$, $F(x) = \int_{g(x)}^{h(x)} f(x,y)\,dy$.

Si f es continua sobre $\{(x,y) \mid a \leqslant x \leqslant b,\ g(x) \leqslant y \leqslant h(x)\}$ y G es cualquier función tal que $D_2\,G(x,y) = f(x,y)$, entonces, según el segundo teorema fundamental del cálculo

$$F(x) = \int_{g(x)}^{h(x)} f(x,y)\,dy = \int_{g(x)}^{h(x)} D_2G(x,y)\,dy = G(x,h(x)) - G(x,g(x)).$$

Pueden también presentarse integrales iteradas en dos dimensiones en que la integración deba efectuarse en primer lugar respecto a x con límites que dependen de y y luego con respecto a y entre límites constantes.

7.2 Ejemplo. Evalúese

$$\int_0^1 \int_{x^2}^{\sqrt{x}} (x^2+y^2)\,dy\,dx.$$

SOLUCIÓN.

$$\begin{aligned}\int_0^1 \int_{x^2}^{\sqrt{x}} (x^2+y^2)\,dy\,dx &= \int_0^1 \int_{x^2}^{\sqrt{x}} D_2(x^2y+\tfrac{1}{3}y^3)\,dy\,dx \\ &= \int_0^1 [x^2y+\tfrac{1}{3}y^3]_{x^2}^{\sqrt{x}}\,dx \\ &= \int_0^1 (x^{5/2}+\tfrac{1}{3}x^{3/2}-x^4-\tfrac{1}{3}x^6)\,dx\end{aligned}$$

$$= [\tfrac{2}{7}x^{7/2} + \tfrac{2}{15}x^{5/2} - \tfrac{1}{5}x^5 - \tfrac{1}{21}x^7]_0^1$$

$$= \tfrac{2}{7} + \tfrac{2}{15} - \tfrac{1}{5} - \tfrac{1}{21} = \tfrac{6}{35}.$$

Problemas

Evalúense las siguientes integrales iteradas.

a) $\displaystyle\int_0^5 \int_{2x}^{x^2} (x+y)\,dy\,dx$ *b*) $\displaystyle\int_2^3 \int_0^{2x} (x^2-y^2)\,dy\,dx$

c) $\displaystyle\int_1^2 \int_y^{2y} x\,dx\,dy$ *d*) $\displaystyle\int_0^1 \int_{\sqrt{y}}^{y} (x^2+y^2)\,dx\,dy$

e) $\displaystyle\int_0^\pi \int_0^{a \operatorname{sen} \theta} r\,dr\,d\theta$ *f*) $\displaystyle\int_0^\pi \int_0^{a \operatorname{sen} \theta} r^2 \cos\theta\,dr\,d\theta$

g) $\displaystyle\int_0^a \int_0^{\sqrt{a^2-x^2}} x\,dy\,dx$ *h*) $\displaystyle\int_0^a \int_0^{\sqrt{a^2-x^2}} y\,dy\,dx$

i) $\displaystyle\int_0^1 \int_0^{x^2} e^{y/x}\,dy\,dx$ *j*) $\displaystyle\int_1^e \int_x^{x^2} \ln y\,dy\,dx.$

8. TEOREMA FUNDAMENTAL PARA LAS INTEGRALES DOBLES

En esta sección probaremos un teorema que establece una relación fundamental entre las integrales dobles y las integrales iteradas. Este teorema nos da un método importante de evaluación de las integrales dobles.

8.1 Teorema. *Si* $\displaystyle\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ *existe y si* $F(x) = \displaystyle\int_{a_2}^{b_2} f(x, y)\,dy$ *existe para cada* $x \in [a_1, b_1]$, *entonces* $\displaystyle\int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x,y)\,dy\,dx = \int_{a_1}^{b_1} F(x)\,dx$ *existe e*

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x,y)\,dy\,dx.$$

PRUEBA. Primero demostraremos que para cualquier partición $P = P_1 \times P_2$ de $[\mathbf{a}, \mathbf{b}]$,

$$L(f, P) \leqslant L(F, P_1) \leqslant U(F, P_1) \leqslant U(f, P).$$

En esta desigualdad $L(f, P)$ es la suma inferior correspondiente a la partición P de $[\mathbf{a}, \mathbf{b}]$ para la integral doble $\displaystyle\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ y $L(F, P_1)$ es la suma inferior correspondiente a la partición P_1 de $[a_1, b_1]$ para la integral simple

$\int_{a_1}^{b_1} F(x)\,dx$. $U(f, P)$ y $U(F, P_1)$ las correspondientes sumas superiores. La existencia de la integral iterada $\int_{a_1}^{b_1} F(x)\,dx = \int_{a_1}^{b_1}\int_{a_2}^{b_2} f(x,y)\,dy\,dx$ y la igualdad de la integral doble $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ y la integral iterada es consecuencia de esta desigualdad.

Sea $P = P_1 \times P_2$ una partición cualquiera de $[\mathbf{a}, \mathbf{b}]$ y sean

$$m_{ij}(f) = \inf\,\{f(x,y)\,|\,x\in[x_{i-1}, x_i],\ y\in[y_{j-1}, y_j]\}$$

$$m_j(f; x) = \inf\,\{f(x,y)\,|\,y\in[y_{j-1}, y_j]\}$$

$$M_{ij}(f) = \sup\,\{f(x,y)\,|\,x\in[x_{i-1}, x_i],\ y\in[y_{j-1}, y_j]\}$$

$$M_j(f; x) = \sup\,\{f(x,y)\,|\,y\in[y_{j-1}, y_j]\}\,.$$

Entonces, para cada $x\in[x_{i-1}, x_i]$,

$$\begin{aligned}\sum_{j=1}^{k_2} m_{ij}(f)\,(y_j - y_{j-1}) &\leqslant \sum_{j=1}^{k_2} m_j(f; x)\,(y_j - y_{j-1}) = L(f, P_2; x)\\ &\leqslant \int_{a_2}^{b_2} f(x,y)\,dy\\ &\leqslant U(f, P_2; x) = \sum_{j=1}^{k_2} M_j(f; x)\,(y_j - y_{j-1})\\ &\leqslant \sum_{j=1}^{k_2} M_{ij}(f)\,(y_j - y_{j-1})\,.\end{aligned}$$

Es decir, para cada $x\in[x_{i-1}, x_i]$,

$$\sum_{j=1}^{k_2} m_{ij}(f)\,(y_j - y_{j-1}) \leqslant F(x) \leqslant \sum_{j=1}^{k_2} M_{ij}(f)\,(y_j - y_{j-1})\,.$$

Como la anterior desigualdad se verifica para todo $x\in[x_{i-1}, x_i]$, concluimos que

8.2 $$\sum_{j=1}^{k_2} m_{ij}(f)\,(y_j - y_{j-1}) \leqslant m_i(F) \leqslant M_i(F) \leqslant \sum_{j=1}^{k_2} M_{ij}(f)\,(y_j - y_{j-1})$$

donde, como habitualmente,

$$m_i(F) = \inf\,\{F(x)\,|\,x\in[x_{i-1}, x_i]\}$$

y

$$M_i(F) = \sup\,\{F(x)\,|\,x\in[x_{i-1}, x_i]\}\,.$$

Multiplicando las desigualdades 8.2 por $x_i - x_{i-1}$ y sumando desde $i = 1$ hasta $i = k_1$, obtenemos

$$\begin{aligned} L(f, P) &= \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} m_{ij}(f)(y_j - y_{j-1})(x_i - x_{i-1}) \\ &\leqslant \sum_{i=1}^{k_1} m_i(F)(x_i - x_{i-1}) = L(F, P_1) \\ &\leqslant U(F, P_1) = \sum_{i=1}^{k_1} M_i(F)(x_i - x_{i-1}) \\ &\leqslant \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} M_{ij}(f)(y_j - y_{j-1})(x_i - x_{i-1}) = U(f, P), \end{aligned}$$

o bien

8.3 $$L(f, P) \leqslant L(F, P_1) \leqslant U(F, P_1) \leqslant U(f, P).$$

Como $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ existe, según el teorema 2.14, pág. 320, dada $\varepsilon > 0$ existe una partición P tal que $U(f, P) - L(f, P) < \varepsilon$. De 8.3 concluimos que

$$U(F, P_1) - L(F, P_1) \leqslant U(f, P) - L(f, P) < \varepsilon$$

y esto implica, por el análogo del teorema 2.14 para integrales simples, que $\int_{a_1}^{b_1} F(x)\,dx$ existe. Además $L(f, P) \leqslant \int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} \leqslant U(f, P)$ y $L(F, P_1) \leqslant \int_{a_1}^{b_1} F(x)\,dx \leqslant U(F, P_1)$ de modo que

$$\left| \int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} - \int_{a_1}^{b_1} F(x)\,dx \right| \leqslant U(f, P) - L(f, P) < \varepsilon.$$

Por tanto

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{a_1}^{b_1} F(x)\,dx = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y)\,dy\,dx.$$

Y esto completa la prueba.

Si intercambiamos x y y en el teorema 8.1, obtenemos

8.4 Teorema. *Si* $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ *existe y* $F(y) = \int_{a_1}^{b_1} f(x, y)\,dx$ *existe para toda* $y \in [a_2, b_2]$, *entonces* $\int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x, y)\,dx\,dy = \int_{a_2}^{b_2} F(y)\,dy$ *existe e*

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x, y)\,dx\,dy.$$

8.5 Ejemplo. Evalúese $\int_{(0,0)}^{(2,5)} (x^2+y^2)\,d\mathbf{x}$.

SOLUCIÓN. Como $\int_{(0,0)}^{(2,5)} (x^2+y^2)\,d\mathbf{x}$ existe y $F(x) = \int_0^5 (x^2+y^2)\,dy$ existe para todo $x\in[0,2]$, por el teorema 8.1,

$$\begin{aligned}\int_{(0,0)}^{(2,5)} (x^2+y^2)\,d\mathbf{x} &= \int_0^2\int_0^5 (x^2+y^2)\,dy\,dx\\ &= \int_0^2 [x^2y+\tfrac{1}{3}y^3]_0^5\,dx\\ &= \int_0^2 [5x^2+\tfrac{125}{3}]\,dx\\ &= \tfrac{5}{3}x^3+\tfrac{125}{3}x]_0^2 = \tfrac{290}{3}.\end{aligned}$$

Probaremos ahora que la integral simple está contenida en la teoría de la integral doble como un caso particular. Supongamos la función F definida y acotada sobre un intervalo $[a, b]$ en R. Definamos la función f sobre el intervalo $[\mathbf{a}, \mathbf{b}]$ en R^2 donde $\mathbf{a} = (a, 0)$ y $\mathbf{b} = (b, 1)$ (figura 9) por la regla de correspondencia:

$$f(\mathbf{x}) = F(x) \text{ para toda } \mathbf{x} = (x, y)\in[\mathbf{a}, \mathbf{b}].$$

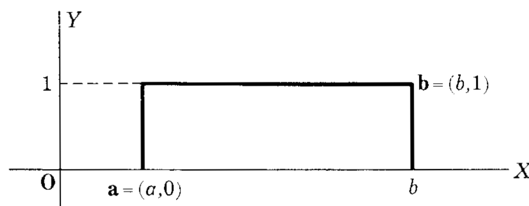


FIGURA 9

Según el teorema 8.1, si f es integrable sobre $[\mathbf{a}, \mathbf{b}]$, entonces

8.6 $$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_a^b\int_0^1 F(x)\,dy\,dx = \int_a^b F(x)\,dx.$$

Nótese que hemos usado el hecho de que $F(x) = \int_0^1 F(x)\,dy$ está definida para toda $x\in[a, b]$.

De los teoremas 8.1 y 5.1 (págs. 347 y 338) obtenemos el siguiente teorema de existencia para las integrales simples como un caso particular de 8.6.

8.7 Teorema. *Si la función real de una variable real F está definida y es acotada sobre* $[a, b]$ *y es continua sobre* $[a, b]$, *excepto en un número finito de puntos, entonces* $\int_a^b F(x)\,dx$ *existe.*

PRUEBA. Sean $x_1, \ldots, x_n$ los puntos de discontinuidad de F donde $a \leqslant x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n \leqslant b$. Definamos f sobre $[\mathbf{a}, \mathbf{b}]$ donde $\mathbf{a} = (a, 0)$ y $\mathbf{b} = (b, 1)$ por la regla de correspondencia $f(\mathbf{x}) = F(\mathbf{x})$ para toda $\mathbf{x} = (x, y) \in [\mathbf{a}, \mathbf{b}]$. La función f es continua sobre $[\mathbf{a}, \mathbf{b}]$ excepto sobre los segmentos rectilíneos desde $(x_i, 0)$ hasta $(x_i, 1)$, $i = 1, \ldots, n$. Como estos segmentos rectilíneos son finitos en número, su unión tiene un área cero (problema 1c, pág. 341). Entonces, por el teorema 5.1, pág. 338, $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ existe. Como $F(x) = \int_0^1 F(x)\,dy$ también existe para todo $x \in [a, b]$, por el teorema 8.1,

$$\int_a^b F(x)\,dx = \int_a^b \int_0^1 F(x)\,dy\,dx = \int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$$

existe.

Damos a continuación condiciones suficientes para asegurar la existencia de las integrales del teorema 8.1.

8.8 Teorema. *Sea f una función acotada sobre un intervalo* $[\mathbf{a}, \mathbf{b}]$ *en* R^2 *y continua sobre* $[\mathbf{a}, \mathbf{b}]$, *excepto sobre un conjunto* $\mathscr{E}$ *de área cero con la propiedad de que toda recta paralela al eje Y intersecta a* $\mathscr{E}$ *en cuando más un número finito de puntos. Entonces*

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y)\,dy\,dx.$$

PRUEBA. Según el teorema 5.1, pág. 338, $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ existe. Además, toda recta paralela al eje Y entre $x = a_1$ y $x = b_1$ intersecta a $\mathscr{E}$ en cuando más un número finito de puntos, de modo que f tiene cuando más un número finito de discontinuidades sobre una tal recta. De donde, según el teorema 8.7,

$$F(x) = \int_{a_2}^{b_2} f(x, y)\,dy$$

existe para todo $x \in [a_1, b_1]$. Luego según el teorema 8.1,

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{a_1}^{b_1} F(x)\,dx = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y)\,dy\,dx.$$

Problemas

1. Encuéntrese $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ para cada una de las siguientes funciones.

a) $f(\mathbf{x}) = x+3y$, $\mathbf{a} = (0, 0)$, $\mathbf{b} = (1, 2)$
b) $f(\mathbf{x}) = x^2 \operatorname{sen} y$, $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2)$
c) $f(\mathbf{x}) = 2x+y-3$, $\mathbf{a} = (1, -1)$, $\mathbf{b} = (4, 2)$
d) $f(\mathbf{x}) = xy^2$, $\mathbf{a} = (0, 0)$, $\mathbf{b} = (2, 3)$

e) $f(\mathbf{x}) = \dfrac{1}{x^3} e^{y/x}$, $\mathbf{a} = (1,0)$, $\mathbf{b} = (3,2)$

f) $f(\mathbf{x}) = y \operatorname{sen} x + x e^y$, $\mathbf{a} = (0, -1)$, $\mathbf{b} = (\pi/2, 1)$.

2. Encuéntrese $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ para cada una de las siguientes funciones.

a) $f(\mathbf{x}) = 3x-y+4$, $[\mathbf{a}, \mathbf{b}]$ acotada por $x = 2$, $x = 0$, $y = -1$, $y = 1$
b) $f(\mathbf{x}) = x+y-xy$, $[\mathbf{a}, \mathbf{b}]$ acotada por $x = 0$, $x = 2$, $y = 0$, $y = 3$

c) $f(\mathbf{x}) = \dfrac{1}{x^2\sqrt{x^2-y^2}}$, $[\mathbf{a}, \mathbf{b}]$ acotada por $x = 1$, $x = 2$, $y = 0$, $y = 1$

d) $f(\mathbf{x}) = \dfrac{1}{x(x^2+y^2)}$, $[\mathbf{a}, \mathbf{b}]$ acotada por $x = 1$, $x = \sqrt{3}$, $y = 0$, $y = 1$.

9. INTEGRALES SOBRE REGIONES EN R^2

En las secciones anteriores consideramos integrales sobre conjuntos acotados $\mathscr{E}$ en R^2. Aquí consideraremos un tipo particular de conjunto $\mathscr{E}$. Pero antes introduciremos alguna terminología.

Recordemos (pág. 187) que un conjunto abierto $\mathscr{E}$ en R^2 se dice que es conexo si no puede representarse como la unión de dos conjuntos abiertos, ajenos y no vacíos. Un conjunto en R^2 se llama *región* si es la unión de un conjunto abierto conexo junto con algunos, ninguno o todos sus puntos frontera. Nosotros usaremos el término región en un sentido más restringido. Supondremos que todas las regiones consideradas en esta y las siguientes secciones son la unión de un número finito de regiones de la forma

$$\mathscr{R}_x = \{(x, y) \mid a \leqslant x \leqslant b,\ g(x) \leqslant y \leqslant h(x)\}$$

o bien

$$\mathscr{R}_y = \{(x, y) \mid a \leqslant y \leqslant b,\ g(y) \leqslant x \leqslant h(y)\}$$

donde g y h son funciones continuas sobre el intervalo cerrado $[a, b]$ en R con $g(x) \leqslant h(x)$ para todo $x \in [a, b]$ (o $g(y) \leqslant h(y)$ para todo $y \in [a, b]$).

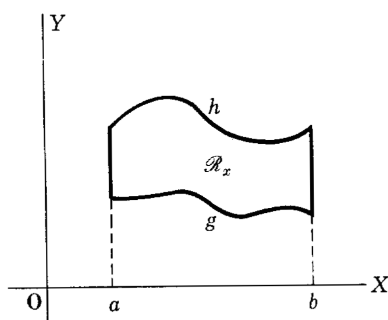


FIGURA 10

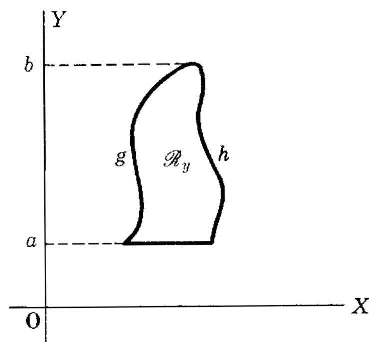


FIGURA 11

Una región del tipo $\mathscr{R}_x$ está ilustrada en la figura 10, y una región del tipo $\mathscr{R}_y$ está ilustrada en la figura 11. En la figura 12, la región $\mathscr{R} = \mathscr{R}_1 \cup \mathscr{R}_2 \cup \mathscr{R}_3$, donde cada una de las regiones $\mathscr{R}_i$ $(i = 1, 2, 3)$ es del tipo $\mathscr{R}_x$.

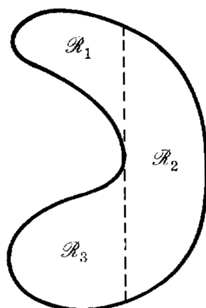
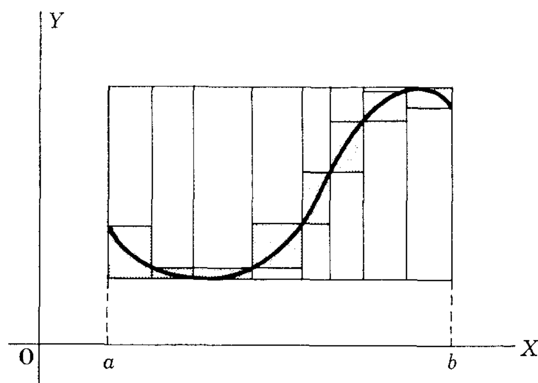


FIGURA 12

Probaremos primero que las regiones (del tipo de las que aquí hemos considerado) tienen área. Por el corolario 5.3 (pág. 339) sabemos que un conjunto acotado $\mathscr{E} \subset \mathrm{R}^2$ tiene área si su frontera tiene área cero. Consideremos una región $\mathscr{R}_x$. Como g y h son continuas sobre $[a, b]$, son acotadas sobre $[a, b]$. Por tanto, $\mathscr{R}_x$ es un conjunto acotado en R^2 y es suficiente demostrar que la frontera de $\mathscr{R}_x$ tiene área cero. Ahora bien, la frontera de $\mathscr{R}_x$ consiste en las gráficas de g y h sobre el intervalo $[a, b]$ y los segmentos rectilíneos desde $(a, g(a))$ hasta $(a, h(a))$ y desde $(b, g(b))$ hasta $(b, h(b))$. Los segmentos rectilíneos tienen área cero. Por ejemplo el segmento entre $(a, g(a))$ y $(a, h(a))$ puede encerrarse en un rectángulo de altura $h(a)-g(a)$ y de ancho $\varepsilon/[h(a)-g(a)+1]$. Este rectángulo tiene área

$$[h(a)-g(a)] \cdot \frac{\varepsilon}{h(a)-g(a)+1} < \varepsilon .$$

FIGURA 13



Así pues, sólo queda por demostrar que la gráfica de una función continua sobre un intervalo $[a, b]$ en R tiene área cero en R^2.

9.1 Teorema. *Si la función f es continua sobre el intervalo cerrado $[a, b]$ en* R, *entonces la gráfica de f tiene área cero en* R^2.

PRUEBA. (Figura 13.) Como f es continua sobre $[a, b]$, f es uniformemente continua sobre $[a, b]$ (teorema 7.6, pág. 478). Luego, dado una $\varepsilon > 0$, podemos encontrar una $\delta > 0$ tal que $|f(x)-f(y)| < \varepsilon$ siempre que $x, y \in [a, b]$ y $|x-y| < \delta$. Sea $P = \{a = x_0 < x_1 < \ldots < x_k = b\}$ una partición de $[a, b]$ de módulo $|P| < \delta$. Como f es continua sobre $[x_{i-1}, x_i]$ para todo $i = 1, \ldots, k$, existen números x_i' y x_i'' en $[x_{i-1}, x_i]$ tales que

$$m_i(f) = f(x_i') \leqslant f(x) \leqslant f(x_i'') = M_i(f)$$

para todo $x \in [x_{i-1}, x_i]$ (teorema 7.9, pág. 479). Así pues, la porción de la gráfica de f entre x_{i-1} y x_i está contenida en el rectángulo $[(x_{i-1}, f(x_i')), (x_i, f(x_i''))]$. Luego si $\bar{A}(f)$ denota el área exterior en R^2 de la gráfica de f, tenemos

$$\bar{A}(f) \leqslant \sum_{i=1}^{k} [f(x_i'')-f(x_i')]\,(x_i - x_{i-1}) < \sum_{i=1}^{k} \varepsilon(x_i - x_{i-1}) = \varepsilon(b-a).$$

Luego, para todo $\varepsilon > 0$, $\bar{A}(f) < \varepsilon(b-a)$; de donde se sigue que la gráfica de f tiene área y que $A(f) = 0$.

9.2 Corolario. *Si $\mathscr{R}$ es una región, entonces $\mathscr{R}$ tiene área.*

PRUEBA. Si $\mathscr{R}$ es una región de tipo $\mathscr{R}_x$, entonces de acuerdo con el teorema 9.1 con f reemplazado por g y h y por la discusión precedente al

teorema 9.1, la frontera de $\mathscr{R}_x$ tiene área cero. Del corolario 5.3 (pág. 339) se sigue que $\mathscr{R}_x$ tiene área. Una región del tipo $\mathscr{R}_y$ se reduce a una región del tipo $\mathscr{R}_x$ si intercambiamos x y y. Si $\mathscr{R}$ es la unión de un número finito de regiones del tipo $\mathscr{R}_x$ y $\mathscr{R}_y$, podemos aplicar el teorema 4.12 (pág. 334) sobre el área de la unión de conjuntos.

Probaremos ahora que si f es continua sobre una región $\mathscr{R}_x$ entonces $\int_{\mathscr{R}_x} f$ puede expresarse en términos de una integral iterada.

9.3 Teorema. *Si f es continua sobre una región $\mathscr{R}_x = \{(x, y) \mid a \leqslant x \leqslant b, g(x) \leqslant y \leqslant h(x)\}$, entonces*

$$\int_{\mathscr{R}_x} f = \int_a^b \int_{g(x)}^{h(x)} f(x, y)\, dy\, dx.$$

PRUEBA. Sean a_2 y b_2 números tales que $a_2 \leqslant g(x)$ y $h(x) \leqslant b_2$ para toda $x \in [a, b]$ y sea $\mathbf{a} = (a, a_2)$, $\mathbf{b} = (b, b_2)$. Entonces $f_{\mathscr{R}_x}$ está acotada sobre $[\mathbf{a}, \mathbf{b}]$ excepto posiblemente sobre las gráficas de g y h que tienen área cero. Por otra parte, cualquier recta paralela al eje Y entre $x = a$ y $x = b$ intersecta las gráficas de g y h sólo una vez cada una de ellas. De donde según el teorema 8.8

$$\begin{aligned}\int_{\mathscr{R}_x} f &= \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{R}_x} = \int_a^b \int_{a_2}^{b_2} f_{\mathscr{R}_x}(x, y)\, dy\, dx \\ &= \int_a^b \int_{a_2}^{g(x)} 0\, dy\, dx + \int_a^b \int_{g(x)}^{h(x)} f(x, y)\, dy\, dx + \int_a^b \int_{h(x)}^{b_2} 0\, dy\, dx \\ &= \int_a^b \int_{g(x)}^{h(x)} f(x, y)\, dy\, dx.\end{aligned}$$

Si intercambiamos x y y en el teorema 9.3, obtenemos

9.4 Corolario. *Si f es continua sobre una región $\mathscr{R}_y = \{(x, y) \mid a \leqslant y \leqslant b, g(y) \leqslant x \leqslant h(y)\}$, entonces*

$$\int_{\mathscr{R}_y} f = \int_a^b \int_{g(y)}^{h(y)} f(x, y)\, dx\, dy.$$

Nota. En la evaluación de una integral doble sobre $\mathscr{R}_x$ por medio de la integral iterada del teorema 9.3 nótese que integramos primero respecto a la y desde la frontera inferior a la frontera superior, y luego respecto a la x sobre la proyección de $\mathscr{R}_x$ sobre el eje X. Para una región $\mathscr{R}_y$, en la integral iterada del corolario 9.4 primero integramos respecto a x desde la frontera izquierda a la frontera derecha y luego con respecto a y sobre la proyección de $\mathscr{R}_y$ sobre el eje Y.

9.5 Ejemplo. Evalúese $\int_{\mathscr{R}} (x^2+y^2)\,d\mathbf{x}$ donde $\mathscr{R}$ es la región en R^2 limitada por las parábolas $y = x^2$ y $x = y^2$ (figura 14).

SOLUCIÓN 1. $\mathscr{R}$ es el conjunto $\mathscr{R} = \mathscr{R}_x = \{(x, y) \mid 0 \leqslant x \leqslant 1,\ x^2 \leqslant y \leqslant \sqrt{x}\}$ de modo que según el teorema 9.3, tenemos

$$\int_{\mathscr{R}} (x^2+y^2)\,d\mathbf{x} = \int_0^1 \int_{x^2}^{\sqrt{x}} (x^2+y^2)\,dy\,dx.$$

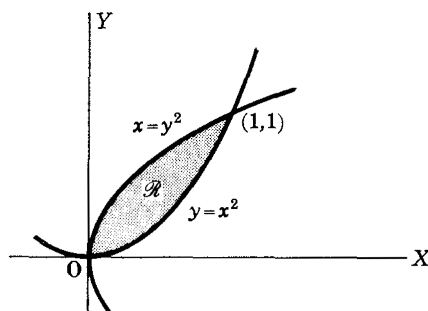


FIGURA 14

Esta integral iterada fue evaluada en el ejemplo 7.2 (pág. 346) donde se encontró que

$$\int_{\mathscr{R}} (x^2+y^2)\,d\mathbf{x} = \tfrac{6}{35}.$$

SOLUCIÓN 2. En este caso $\mathscr{R}$ también es una región del tipo $\mathscr{R}_y$: $\mathscr{R} = \mathscr{R}_y = \{(x, y) \mid 0 \leqslant y \leqslant 1,\ y^2 \leqslant x \leqslant \sqrt{y}\}$. De donde, según el corolario 9.4, tenemos

$$\int_{\mathscr{R}} (x^2+y^2)\,d\mathbf{x} = \int_0^1 \int_{y^2}^{\sqrt{y}} (x^2+y^2)\,dx\,dy = \tfrac{6}{35}.$$

Supongamos que $\mathscr{R}$ puede expresarse como la unión de un conjunto finito de regiones ajenas $\{\mathscr{R}_i \mid i = 1, \ldots, n\}$ del tipo $\mathscr{R}_x$ y $\mathscr{R}_y$, es decir,

$$\mathscr{R} = \bigcup_{i=1}^{n} \mathscr{R}_i \quad \text{donde} \quad A(\mathscr{R}_i \cap \mathscr{R}_j) = 0 \ \text{para} \ i \neq j.$$

Si f es continua sobre $\mathscr{R}$, como cada una de las $\mathscr{R}_i$ tiene área, según el corolario 5.6, pág. 340, $\int_{\mathscr{R}_i} f$ existe. Luego conforme al corolario 6.13, pág. 345,

$$\int_{\mathscr{R}} f = \sum_{i=1}^{n} \int_{\mathscr{R}_i} f.$$

Problemas

Encuéntrese $\int_{\mathscr{R}} f$ para las siguientes funciones si $\mathscr{R}$ es la región limitada por las curvas dadas.

a) $f(x, y) = y$; $\mathscr{R}$ limitada a la izquierda por $y^2 = x$, y a la derecha por $x^2+y^2 = 2$.

b) $f(x, y) = x$; $\mathscr{R}$ limitada por $y^2 = x$ y $x^2 = y$.

c) $f(x, y) = y^2$; $\mathscr{R}$ limitada por $y = \operatorname{sen} x$, $y = \cos x$, $x \in \left[0, \frac{\pi}{4}\right]$.

d) $f(x, y) = x+2y-5$; $\mathscr{R}$ limitada por $y^2 = x$ y $x^2 = y$.

e) $f(x, y) = \sqrt{1-x^2-y^2}$; $\mathscr{R}$ limitada por una circunferencia de radio 1 y centro en el origen.

f) $f(x, y) = y$; $\mathscr{R}$ limitada por $x^2+y^2 = 2$ por arriba, y por abajo por $x^2 = y$.

g) $f(x, y) = e^{y/x}$; $\mathscr{R}$ limitada por $x = 0$, $x = 1$, $y = 0$, $y = x^2$.

h) $f(x, y) = y^2 \operatorname{sen} x$; $\mathscr{R}$ limitada por $x = 0$, $x = \pi$, $y = 0$, $y = 1 + \cos x$.

i) $f(x, y) = x+y$; $\mathscr{R}$ limitada por $x^2+y^2 = a^2$.

j) $f(x, y) = \operatorname{senh} \frac{y}{x} \cosh x$; $\mathscr{R}$ limitada por $x = 0$, $x = 1$, $y = 2x$, $y = 2$.

k) $f(x, y) = 4-x^2$; $\mathscr{R}$ limitada por $2x+y = 6$, $y = x$, $x = 0$.

l) $f(x, y) = 2-2y^2-\frac{1}{2}x^2$; $\mathscr{R}$ limitada por $x+2y = 2$, $x = 0$, $y = 0$.

10. ÁREA Y MOMENTOS DE REGIONES PLANAS

La determinación del área de las regiones del tipo $\mathscr{R}_x$ se considera usualmente en el cálculo de funciones reales de una variable real. Si

$$\mathscr{R}_x = \{(x, y) \mid a \leqslant x \leqslant b,\ g(x) \leqslant y \leqslant h(x)\},$$

entonces el área de $\mathscr{R}_x$ es

$$A(\mathscr{R}_x) = \int_a^b [h(x)-g(x)]\, dx\,.$$

De acuerdo con el teorema 4.6, pág. 332, tenemos

$$A(\mathscr{R}_x) = \int_{\mathscr{R}_x} 1\,.$$

Por el teorema 9.3,

$$\int_{\mathscr{R}_x} 1 = \int_a^b \int_{g(x)}^{h(x)} dy\, dx = \int_a^b [h(x)-g(x)]\, dx\,.$$

Así pues, para regiones del tipo $\mathscr{R}_x$ estos dos métodos de determinación del área son acordes y esencialmente el mismo.

10.1 Ejemplo. Encuéntrese el área de la región limitada por la parábola $y = x^2+2$ y la recta $y = x+4$.

SOLUCIÓN. (Figura 15.) Los puntos de intersección de la parábola y la recta se encuentra que son $(-1, 3)$ y $(2, 6)$ y la región

$$\mathscr{R}_x = \{(x, y) \mid -1 \leqslant x \leqslant 2,\ x^2+2 \leqslant y \leqslant x+4\}.$$

Entonces

$$A(\mathscr{R}_x) = \int_{\mathscr{R}_x} 1 = \int_{-1}^{2} \int_{x^2+2}^{x+4} dy\,dx = \int_{-1}^{2} (x+4-x^2-2)\,dx$$

$$= \frac{x^2}{2} + 2x - \frac{x^3}{3}\Big]_{-1}^{2} = \frac{9}{2}.$$

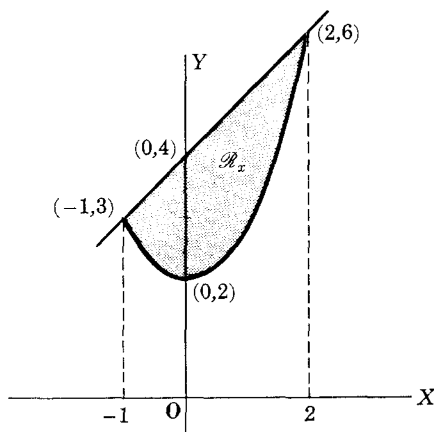


FIGURA 15

Definimos ahora el primer momento y el segundo momento de una región plana con respecto a una recta.

10.2 Definición. (Figura 16.) *Sea $\mathscr{L}$ una recta con normal* $\mathbf{n}$ *y sea* $\mathbf{x}_0$ *un punto fijo de $\mathscr{L}$. El (primer)* ***momento*** *$M_{\mathscr{L}}$ y el* ***momento de inercia*** *(segundo momento) $I_{\mathscr{L}}$ de una región $\mathscr{R}$ en* R^2 *con respecto a la recta $\mathscr{L}$ se definen por*

$$M_{\mathscr{L}} = \int_{\mathscr{R}} \mathrm{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0)\,d\mathbf{x}$$

e

$$I_{\mathscr{L}} = \int_{\mathscr{R}} [\text{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0)]^2 \, d\mathbf{x}$$

respectivamente.

El número $\text{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0) = \dfrac{\mathbf{n}\cdot(\mathbf{x}-\mathbf{x}_0)}{|\mathbf{n}|}$ es la distancia dirigida del punto $\mathbf{x}$ a la recta $\mathscr{L}$. Así pues, $M_{\mathscr{L}}$ es la integral sobre $\mathscr{R}$ de la distancia dirigida de $\mathscr{L}$ a los puntos de $\mathscr{R}$ *e* $I_{\mathscr{L}}$ es la integral sobre $\mathscr{R}$ del cuadrado de la distancia de $\mathscr{L}$ a los puntos de $\mathscr{R}$. El signo de $M_{\mathscr{L}}$ depende de la elección de $\mathbf{n}$ ya que $\text{Comp}_{-\mathbf{n}}(\mathbf{x}-\mathbf{x}_0) = -\text{Comp}\ (\mathbf{x}-\mathbf{x}_0)$.

Sea $ax+by+c = 0$ una ecuación de la recta $\mathscr{L}$. El vector $\mathbf{n} = (a, b)$ es una normal a $\mathscr{L}$. Si $\mathbf{x}_0 = (x_0, y_0)$ es un punto de $\mathscr{L}$ y $\mathbf{x} = (x, y)$ es un punto de $\mathscr{R}$, entonces

$$\mathbf{n}\cdot\mathbf{x}_0 = ax_0+by_0 = -c$$

y

$$\begin{aligned} \text{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0) &= \frac{\mathbf{n}\cdot(\mathbf{x}-\mathbf{x}_0)}{|\mathbf{n}|} = \frac{\mathbf{n}\cdot\mathbf{x}-\mathbf{n}\cdot\mathbf{x}_0}{|\mathbf{n}|} \\ &= \frac{1}{\sqrt{a^2+b^2}}(ax+by+c). \end{aligned} \tag{10.3}$$

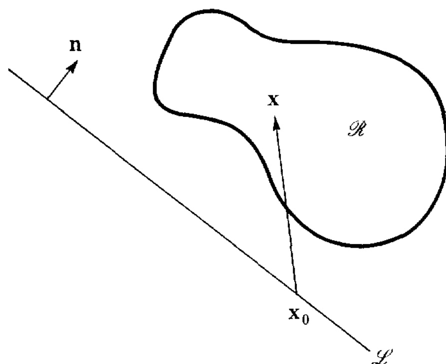


FIGURA 16

Así pues, el primer momento y el segundo momento pueden expresarse de las formas

$$M_{\mathscr{L}} = \frac{1}{\sqrt{a^2+b^2}} \int_{\mathscr{R}} (ax+by+c)\, d\mathbf{x}$$

e

$$I_{\mathscr{L}} = \frac{1}{a^2+b^2}\int_{\mathscr{R}} (ax+by+c)^2\,d\mathbf{x}\,.$$

10.4 Ejemplo. Encuéntrense los momentos primero y segundo con respecto a la recta $\mathscr{L}$: $ax+by+c = 0$ de la región $\mathscr{R}_x$ limitada por la recta $y = x+1$ y la parábola $y = (x-1)^2$.

SOLUCIÓN. (Figura 17.)

$$M_{\mathscr{L}} = \int_{\mathscr{R}_x} \operatorname{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0)\,d\mathbf{x} = \frac{1}{\sqrt{a^2+b^2}}\int_0^3\int_{(x-1)^2}^{x+1}(ax+by+c)\,dy\,dx$$

$$= \frac{1}{\sqrt{a^2+b^2}}\int_0^3\left[axy+\frac{1}{2}by^2+cy\right]_{(x-1)^2}^{x+1}dx$$

$$= \frac{1}{\sqrt{a^2+b^2}}\int_0^3\{a[-x^3+3x^2]+\frac{1}{2}b[(x+1)^2-(x-1)^4]$$

$$+c[-x^2+3x]\}\,dx$$

$$= \frac{1}{\sqrt{a^2+b^2}}\left\{\frac{27}{4}a+\frac{36}{5}b+\frac{9}{2}c\right\}.$$

$$I_{\mathscr{L}} = \int_{\mathscr{R}_x}[\operatorname{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0)]^2\,d\mathbf{x} = \frac{1}{a^2+b^2}\int_0^3\int_{(x-1)^2}^{x+1}(ax+by+c)^2\,dy\,dx$$

$$= \frac{1}{a^2+b^2}\left[\frac{243}{20}a^2+\frac{423}{28}b^2+\frac{9}{2}c^2+\frac{513}{20}ab+\frac{27}{2}ac+\frac{72}{5}bc\right].$$

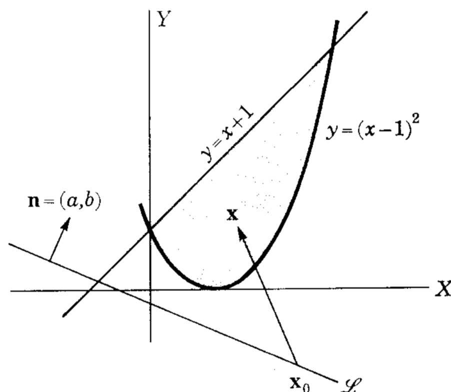


FIGURA 17

En la definición de $M_{\mathscr{L}}$ y en la de $I_{\mathscr{L}}$ cuando $\mathscr{L}$ es el eje X, es habitual escoger $\mathbf{n}$ en la dirección positiva del eje Y. Entonces

$$\text{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0) = \mathbf{j}\cdot(\mathbf{x}-\mathbf{x}_0) = y$$

y

$$M_x = \int_{\mathscr{R}} y\,d\mathbf{x}, \qquad I_x = \int_{\mathscr{R}} y^2\,d\mathbf{x}.$$

Cuando $\mathscr{L}$ es el eje Y, lo habitual es escoger $\mathbf{n}$ en la dirección positiva del eje X. Entonces

$$\text{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0) = \mathbf{i}\cdot(\mathbf{x}-\mathbf{x}_0) = x$$

y

$$M_y = \int_{\mathscr{R}} x\,d\mathbf{x}, \qquad I_y = \int_{\mathscr{R}} x^2\,d\mathbf{x}.$$

Además del momento de inercia con respecto a rectas, se define también el momento de inercia con respecto al origen. A éste se le llama momento polar de inercia y está definido por:

$$J = I_z = \int_{\mathscr{R}} |\mathbf{x}|^2\,d\mathbf{x} = \int_{\mathscr{R}} (x^2+y^2)\,d\mathbf{x} = I_y+I_x.$$

10.5 Ejemplo. Encuéntrense los momentos primero y segundo con respecto a los ejes X y Y y el momento polar de inercia de la región limitada por la parábola $y = x^2+2$ y la recta $y = x+4$.

SOLUCIÓN. La región es la misma que la del ejemplo 10.1 y aparece en la figura 15. $\mathscr{R} = \{(x, y) \mid -1 \leqslant x \leqslant 2,\, x^2+2 \leqslant y \leqslant x+4\}$.

$$M_x = \int_{\mathscr{R}} y\,d\mathbf{x} = \int_{-1}^{2} \int_{x^2+2}^{x+4} y\,dy\,dx = 81/5$$

$$M_y = \int_{\mathscr{R}} x\,d\mathbf{x} = \int_{-1}^{2} \int_{x^2+2}^{x+4} x\,dy\,dx = 9/4$$

$$I_x = \int_{\mathscr{R}} y^2\,d\mathbf{x} = \int_{-1}^{2} \int_{x^2+2}^{x+4} y^2\,dy\,dx = 8667/140$$

$$I_y = \int_{\mathscr{R}} x^2\,d\mathbf{x} = \int_{-1}^{2} \int_{x^2+2}^{x+4} x^2\,dy\,dx = 63/20$$

$$I_z = I_x+I_y = 2277/35.$$

10.6 Definición. *El punto*

$$\bar{\mathbf{x}} = (\bar{x}, \bar{y}) = \left(\frac{M_y}{A(\mathscr{R})}, \frac{M_x}{A(\mathscr{R})}\right)$$

se llama ***centroide*** *de una región* $\mathscr{R}$.

10.7 Teorema. *Si* $\mathscr{L}$ *es una recta que pasa por el centroide de una región plana* $\mathscr{R}$, *entonces* $M_{\mathscr{L}} = 0$.

PRUEBA. Sea $\mathbf{n} = (n_1, n_2)$ una normal unitaria a $\mathscr{L}$. Entonces, como $\bar{\mathbf{x}}$ es un punto sobre $\mathscr{L}$, tenemos

$$\begin{aligned} M_{\mathscr{L}} &= \int_{\mathscr{R}} \operatorname{Comp}_{\mathbf{n}}(\mathbf{x} - \bar{\mathbf{x}})\, d\mathbf{x} = \int_{\mathscr{R}} \mathbf{n} \cdot (\mathbf{x} - \bar{\mathbf{x}})\, d\mathbf{x} \\ &= \int_{\mathscr{R}} [n_1(x - \bar{x}) + n_2(y - \bar{y})]\, d\mathbf{x} \\ &= n_1(M_y - \bar{x}A(\mathscr{R})) + n_2(M_x - \bar{y}A(\mathscr{R})) = 0. \end{aligned}$$

10.8 Teorema. *Si* $\mathscr{L}_1$ *y* $\mathscr{L}_2$ *son rectas paralelas de normal* $\mathbf{n}$, $\mathbf{x}_1$ *y* $\mathbf{x}_2$ *puntos fijos de* $\mathscr{L}_1$ *y* $\mathscr{L}_2$ *respectivamente, y* $\mathscr{R}$ *es una región, entonces*

$$\begin{aligned} I_{\mathscr{L}_1} &= I_{\mathscr{L}_2} + 2[\operatorname{Comp}_{\mathbf{n}}(\mathbf{x}_2 - \mathbf{x}_1)]M_{\mathscr{L}_2} + [\operatorname{Comp}_{\mathbf{n}}(\mathbf{x}_2 - \mathbf{x}_1)]^2 A(\mathscr{R}) \\ &= I_{\mathscr{L}_2} + 2dM_{\mathscr{L}_2} + d^2 A(\mathscr{R}) \end{aligned}$$

donde $A(\mathscr{R})$ *es el área de* $\mathscr{R}$ *y* $d = \operatorname{Comp}_{\mathbf{n}}(\mathbf{x}_2 - \mathbf{x}_1)$ *es la distancia dirigida de* $\mathscr{L}_1$ *a* $\mathscr{L}_2$.

PRUEBA. Como $\mathbf{x} - \mathbf{x}_1 = (\mathbf{x} - \mathbf{x}_2) + (\mathbf{x}_2 - \mathbf{x}_1)$, tenemos

$$\begin{aligned} I_{\mathscr{L}_1} &= \int_{\mathscr{R}} [\operatorname{Comp}_{\mathbf{n}}(\mathbf{x} - \mathbf{x}_1)]^2\, d\mathbf{x} = \int_{\mathscr{R}} \left[\frac{\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_1)}{|\mathbf{n}|}\right]^2 d\mathbf{x} \\ &= \int_{\mathscr{R}} \left[\frac{\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_2)}{|\mathbf{n}|} + \frac{\mathbf{n} \cdot (\mathbf{x}_2 - \mathbf{x}_1)}{|\mathbf{n}|}\right]^2 d\mathbf{x} \\ &= \int_{\mathscr{R}} [\operatorname{Comp}_{\mathbf{n}}(\mathbf{x} - \mathbf{x}_2) + \operatorname{Comp}_{\mathbf{n}}(\mathbf{x}_2 - \mathbf{x}_1)]^2\, d\mathbf{x} \\ &= \int_{\mathscr{R}} [\operatorname{Comp}_{\mathbf{n}}(\mathbf{x} - \mathbf{x}_2)]^2\, d\mathbf{x} + 2 \operatorname{Comp}_{\mathbf{n}}(\mathbf{x}_2 - \mathbf{x}_1) \int_{\mathscr{R}} \operatorname{Comp}_{\mathbf{n}}(\mathbf{x} - \mathbf{x}_2)\, d\mathbf{x} \\ &\quad + [\operatorname{Comp}_{\mathbf{n}}(\mathbf{x}_2 - \mathbf{x}_1)]^2 \int_{\mathscr{R}} d\mathbf{x} \\ &= I_{\mathscr{L}_2} + 2dM_{\mathscr{L}_2} + d^2 A(\mathscr{R}). \end{aligned}$$

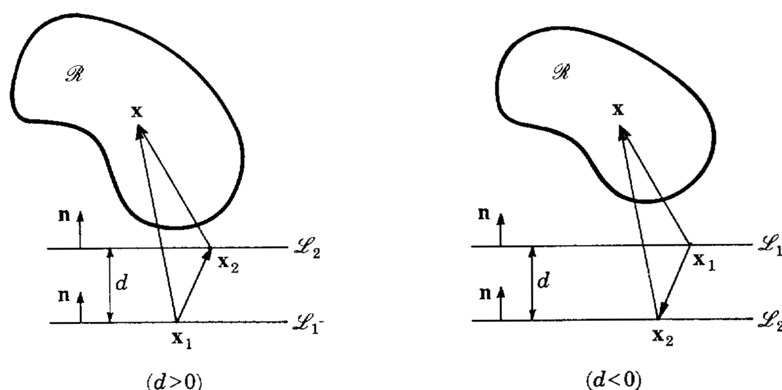


FIGURA 18

Un resultado análogo para los primeros momentos se da en el problema 9.

10.9 Corolario. (Teorema de los ejes paralelos.) *Si $\mathscr{L}_2$ es una recta que pasa por el centroide de $\mathscr{R}$ y $\mathscr{L}_1$ es una recta paralela a $\mathscr{L}_2$, entonces*

$$I_{\mathscr{L}_1} = I_{\mathscr{L}_2} + d^2 A(\mathscr{R}).$$

PRUEBA. Según el teorema 10.7, $M_{\mathscr{L}_2} = 0$.

Problemas

1. Encuéntrense los centroides de las regiones limitadas por los siguientes conjuntos de curvas:

a) $x^{1/2}+y^{1/2} = a^{1/2}$ y los ejes coordenados
b) la hipérbola $y = 4/(4-x)$ y la parábola $y = (x-1)^2$
c) por la parte superior por la elipse $25x^2+16y^2 = 400$ y por la parte inferior por el eje X
d) la región del primer cuadrante limitada por $y = x^3$ y $x = y^3$
e) $y^2 = x^3$ y $y = x$
f) la región del primer cuadrante en el interior de la elipse $\dfrac{x^2}{4} + \dfrac{y^2}{9} = 1$ y fuera de la elipse $x^2 + \dfrac{y^2}{9} = 1$.

2. Encuéntrense los momentos de inercia con respecto a cada uno de los ejes coordenados de las regiones limitadas por los siguientes conjuntos de curvas:

a) $y = 2x-x^2$ y $y = x$
b) $y = 2x-x^2$ y $x+y = 2$

c) a la izquierda por $y^2 = 4x$, a la derecha por $x^2+y^2 = 32$ y en la parte inferior por el eje X

d) $y^2 = 1+x$ y $y^2 = 1-2x$

e) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$

f) $\dfrac{x}{a} + \dfrac{y}{b} = 1$ y los ejes coordenados.

3. Encuéntrese el área, M_x, M_y, I_x, I_y y el centroide de cada una de las siguientes regiones:

a) limitada en la parte de arriba por $x^2+y^2 = 2$ y en la parte de abajo por $y = x^2$

b) limitada a la izquierda por $y = x^2$, a la derecha por $x^2+y^2 = 2$ y debajo por el eje X

c) limitada por $y = x^2$ y $x = y^2$

d) limitada por $y = \operatorname{sen} x$, $y = \cos x$, $x \in [-\pi/4, \pi/4]$

e) limitada por $y = x^2+2$ y $y = x+4$

f) limitada en la parte superior por $x^2+y^2 = a^2$ y en la parte inferior por el eje X.

4. Sea $\mathscr{R}$ la región limitada por $y = x^2$ y $y^2 = x$. Encuéntrese:

a) $M_{\mathscr{L}}$ cuando $\mathscr{L}$ es la recta $y = 1$

b) $M_{\mathscr{L}}$ cuando $\mathscr{L}$ es la recta $x = -1$

c) $M_{\mathscr{L}}$ cuando $\mathscr{L}$ es la recta $x+y = 0$

d) $M_{\mathscr{L}}$ cuando $\mathscr{L}$ es la recta $x-y = 0$

e) $I_{\mathscr{L}}$ cuando $\mathscr{L}$ es la recta $y = -3$

f) $I_{\mathscr{L}}$ cuando $\mathscr{L}$ es la recta $x = 2$

g) $I_{\mathscr{L}}$ cuando $\mathscr{L}$ es la recta $x+y+5 = 0$

h) $I_{\mathscr{L}}$ cuando $\mathscr{L}$ es la recta $x-y = 0$.

5. Demuéstrese que el centroide de un triángulo está a un tercio de la distancia de un lado al vértice opuesto.

6. Demuéstrese que si una región $\mathscr{R}$ tiene un eje de simetría, entonces el centroide se encuentra sobre este eje de simetría.

7. Demuéstrese que el área de una región limitada por una parábola y una cuerda ortogonal al eje de la parábola es igual a dos tercios del área de un rectángulo circunscrito.

8. Encuéntrese el centroide de la región del problema 7 y los momentos de inercia de esta región con respecto a los ejes que pasan por el centroide paralelo y perpendicular al eje de la parábola.

9. Pruébese que si $\mathscr{L}_1$ y $\mathscr{L}_2$ son rectas paralelas de normal $\mathbf{n}$, $\mathbf{x}_1$ y $\mathbf{x}_2$ puntos fijos sobre $\mathscr{L}_1$ y $\mathscr{L}_2$ respectivamente, y $\mathscr{R}$ es una región, entonces

$$M_{\mathscr{L}_1} = M_{\mathscr{L}_2} + [\text{Comp}_{\mathbf{n}}(\mathbf{x}_2 - \mathbf{x}_1)]\, A(\mathscr{R}).$$

10. Pruébese que si los ejes de coordenadas son elegidos de modo que el origen esté en el centroide de una región $\mathscr{R}$, entonces para una recta cualquiera $\mathscr{L}$ que pase por el centroide con ángulo de inclinación α,

$$I_{\mathscr{L}} = I_y \operatorname{sen}^2 \alpha - 2I_{xy} \operatorname{sen} \alpha \cos \alpha + I_x \cos^2 \alpha$$

donde $I_{xy} = \int_{\mathscr{R}} xy\, d\mathbf{x}$. Al número I_{xy} se le llama producto de inercia.

Sugerencia. Tómese el punto $\mathbf{x}_0 = \mathbf{0}$ sobre $\mathscr{L}$ y $\mathbf{n} = (\operatorname{sen} \alpha, -\cos \alpha)$.

11. Encuéntrese el momento de inercia de un rectángulo de base w y altura h con respecto a la recta $\mathscr{L}$: $ax+by+c = 0$. Especialícese el resultado para obtener el momento de inercia con respecto a:

a) la base
b) una recta que pasa por el centroide paralela a la base
c) una diagonal.

11. VOLUMEN BAJO UNA SUPERFICIE

En esta sección consideraremos el volumen de una región en R^3 de un tipo muy especial. Suponemos que el volumen tiene propiedades similares a las del área. Si $\mathscr{E}$ es un conjunto de puntos en R^3 con volumen, denotaremos el volumen de $\mathscr{E}$ por $V(\mathscr{E})$. Para aquellos conjuntos de puntos $\mathscr{E}$, $\mathscr{F}$, $\mathscr{G}$, ... para los que el volumen está definido suponemos

11.1 $V(\mathscr{E}) \geqslant 0$.

11.2 *Si* $\mathscr{E} \subset \mathscr{F}$, *entonces* $V(\mathscr{E}) \leqslant V(\mathscr{F})$.

11.3 *Si* $\mathscr{G} = \mathscr{E} \cup \mathscr{F}$ *y* $V(\mathscr{E} \cap \mathscr{F}) = 0$, *entonces* $V(\mathscr{G}) = V(\mathscr{E}) + V(\mathscr{F})$.

11.4 *El volumen de un paralelepípedo rectangular es el producto de las longitudes de los lados.*

Supongamos que una región en R^3 está lateralmente limitada por una superficie cilíndrica con un generador paralelo al eje Z, limitada superiormente por una superficie $z = f(x, y)$, y limitada inferiormente por una región cerrada y acotada $\mathscr{R}$ del plano XY (figura 19). Sea $[\mathbf{a}, \mathbf{b}]$ un intervalo en R^2 tal que $\mathscr{R} \subset [\mathbf{a}, \mathbf{b}]$ y sea P una partición de $[\mathbf{a}, \mathbf{b}]$ en subintervalos $\mathscr{R}_i$.

Definamos

$$m_i(f_{\mathscr{R}}) = \inf \{f_{\mathscr{R}}(x, y) \mid (x, y) \in \mathscr{R}_i\}$$

$$M_i(f_{\mathscr{R}}) = \sup \{f_{\mathscr{R}}(x, y) \mid (x, y) \in \mathscr{R}_i\}.$$

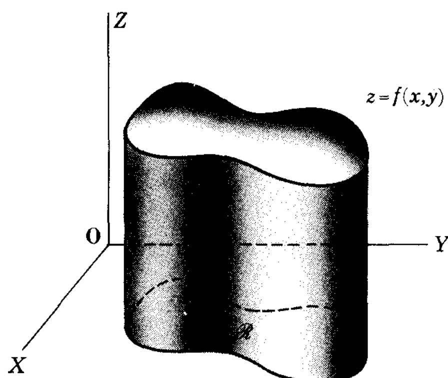


FIGURA 19

Entonces, el volumen V_i del cilindro de base $\mathscr{R}_i$ acotado superiormente por $z = f_{\mathscr{R}}(x, y)$ es una cantidad intermedia entre $m_i(f_{\mathscr{R}})\, A(\mathscr{R}_i)$ y $M_i(f_{\mathscr{R}})\, A(\mathscr{R}_i)$ donde $A(\mathscr{R}_i)$ es el área del subintervalo $\mathscr{R}_i$:

$$m_i(f_{\mathscr{R}})\, A(\mathscr{R}_i) \leqslant V_i \leqslant M_i(f_{\mathscr{R}})\, A(\mathscr{R}_i).$$

De donde

$$L(f_{\mathscr{R}}, P) = \sum_{i=1}^{k} m_i(f_{\mathscr{R}}) A(\mathscr{R}_i) \leqslant V \leqslant \sum_{i=1}^{k} M_i(f_{\mathscr{R}}) A(\mathscr{R}_i) = U(f_{\mathscr{R}}, P).$$

Por tanto, si f está acotado sobre $\mathscr{R}$,

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f_{\mathscr{R}} = \underline{\int_{\mathscr{R}}} f \leqslant V \leqslant \overline{\int_{\mathscr{R}}} f = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f_{\mathscr{R}}$$

y f es integrable sobre $\mathscr{R}$

$$V = \int_{\mathscr{R}} f.$$

Por el teorema 9.3, pág. 355, si $\mathscr{R} = \mathscr{R}_x = \{(x, y) \mid a \leqslant x \leqslant b, g(x) \leqslant y \leqslant h(x)\}$ y f es continua sobre $\mathscr{R}$, luego

11.5 $$V = \int_{\mathscr{R}} f = \int_a^b \int_{g(x)}^{h(x)} f(x, y)\, dy\, dx$$

y por el corolario 9.4 si $\mathscr{R} = \mathscr{R}_y = \{(x, y) \mid a \leqslant y \leqslant b, g(y) \leqslant x \leqslant h(y)\}$ y f es continua sobre $\mathscr{R}$, entonces

11.6 $$V = \int_{\mathscr{R}} f = \int_a^b \int_{g(y)}^{h(y)} f(x, y)\, dx\, dy.$$

11.7 Ejemplo. Encuéntrese el volumen limitado en la parte superior por el paraboloide de revolución $z = x^2 + y^2$, inferiormente por el plano XY, y lateralmente por el cilindro circular $x^2 + y^2 = 4$.

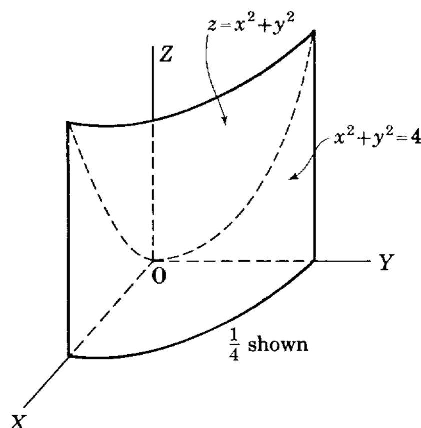


FIGURA 20

Solución. (Figura 20.) La región $\mathscr{R}$ en el plano XY sobre la que integramos es la base del cilindro:

$$\mathscr{R} = \mathscr{R}_x = \{(x, y) \mid -2 \leqslant x \leqslant 2,\ -\sqrt{4-x^2} \leqslant y \leqslant \sqrt{4-x^2}\}.$$

De donde

$$V = \int_{\mathscr{R}} (x^2 + y^2)\, d\mathbf{x} = \int_{-2}^{2} \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} (x^2 + y^2)\, dy\, dx = 8\pi.$$

Como tanto el integrando como la región $\mathscr{R}$ son simétricos con respecto a los planos XZ y YZ, podemos escribir

$$V = 4 \int_0^2 \int_0^{\sqrt{4-x^2}} (x^2 + y^2)\, dy\, dx.$$

Podíamos haber descrito nuestra región $\mathscr{R}$ como

$$\mathscr{R} = \mathscr{R}_y = \{(x, y) \mid -2 \leqslant y \leqslant 2,\ -\sqrt{4-y^2} \leqslant x \leqslant \sqrt{4-y^2}\}$$

y en este caso habríamos obtenido

$$V = \int_{-2}^{2} \int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} (x^2+y^2)\,dx\,dy = 8\pi .$$

11.8 Ejemplo. Encuéntrese el volumen del sólido limitado por el plano $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$ (a, b, c todos positivos) y los planos coordenados.

Solución. (Figura 21.) La región base $\mathscr{R}$ es

$$\mathscr{R}_x = \{(x, y) \mid 0 \leqslant x \leqslant a,\ 0 \leqslant y \leqslant \frac{b}{a}(a-x)\} .$$

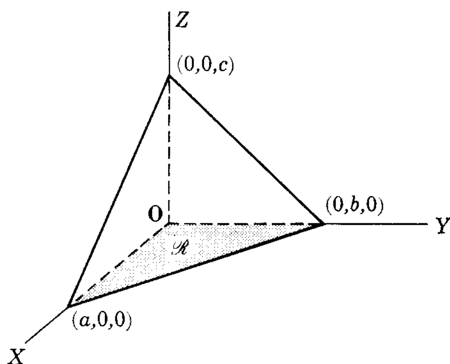


FIGURA 21

De donde

$$V = \int_{\mathscr{R}_x} c\left(1 - \frac{x}{a} - \frac{y}{b}\right) d\mathbf{x} = \int_0^a \int_0^{\frac{b}{a}(a-x)} c\left(1 - \frac{x}{a} - \frac{y}{b}\right) dy\,dx = \frac{abc}{6}.$$

Problemas

1. Encuéntrese el volumen del sólido que está sobre la región $\mathscr{R}$ y tiene la gráfica de f como su superficie superior. Dibújese una figura.

a) $\mathscr{R}$ es el triángulo limitado por los ejes de coordenadas y la recta $x+y = 1$; $f(x, y) = 1-x^2$.

b) $\mathscr{R}$ es el rectángulo limitado por los ejes de coordenadas y las rectas $x = 2$, $y = 3$; $f(x, y) = x^2+y^2$.

c) $\mathscr{R}$ está limitada por la circunferencia $x^2+y^2 = 1$; $f(x, y) = y+2$.

d) $\mathscr{R}$ está limitada por la circunferencia $x^2+y^2 = 1$; $f(x, y) = 3x+4$.

2. Encuéntrese el volumen de un hemisferio de radio la unidad.

3. Encuéntrese el volumen de la región en R^3 limitada superiormente por $z = 1-x^2-y^2$ y debajo por el plano XY.

4. Encuéntrese el volumen de la región en R^3 limitada arriba por $z = 4-x$, lateralmente por $x^2+y^2 = 1$, y debajo por el plano XY.

12. VOLÚMENES DE REVOLUCIÓN Y EL TEOREMA DE PAPPUS

En la sección 11 consideramos el volumen de un tipo muy particular de región en R^3 —regiones limitadas lateralmente por una superficie cilíndrica con un generador paralelo al eje Z, arriba por una superficie $z = f(x, y)$ y debajo por el plano XY. En esta sección consideraremos el volumen de un sólido de revolución generado al hacer girar una región plana alrededor de una recta del plano.

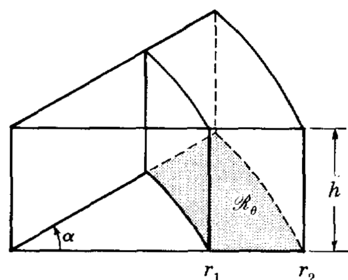


FIGURA 22

La región en R^3 (figura 22) que se encuentra sobre $\mathscr{R}_\theta = \{(r\cos\theta, r\operatorname{sen}\theta) \mid 0 \leqslant \theta < \alpha, r_1 \leqslant r < r_2\}$ y está limitada superiormente por el plano $z = h$ tiene volumen $\alpha\bar{r}(r_2-r_1)h$ donde $\bar{r} = \frac{1}{2}(r_1+r_2)$. Esta región puede imaginarse que está generada por el giro de un rectángulo de lados r_2-r_1 y h un ángulo α alrededor de una recta paralela al lado de longitud h a una distancia $\frac{1}{2}(r_1+r_2)$ del centroide del rectángulo.

Supongamos que $\mathscr{R}$ es una región en el plano (figura 23). Sea $[\mathbf{a}, \mathbf{b}]$ un intervalo tal que $\mathscr{R} \subset [\mathbf{a}, \mathbf{b}]$ y sea P una partición de $[\mathbf{a}, \mathbf{b}]$ en subintervalos $\mathscr{R}_{ij}$. Si el intervalo $[\mathbf{a}, \mathbf{b}]$ se hace girar un ángulo α alrededor de la recta $y = c$ donde $c \leqslant a_2$ o $c \geqslant b_2$, entonces el subintervalo $\mathscr{R}_{ij}$ genera una región en R^3 de volumen: $\alpha|\bar{y}_j - c|\, A(\mathscr{R}_{ij})$ donde $\bar{y}_j = \frac{1}{2}(y_{j-1}+y_j)$. Supongamos que la intersección de esta región con el sólido de revolución

generado haciendo girar $\mathscr{R}$ un ángulo α alrededor de $y = c$ tiene un volumen V_{ij}. Entonces

$$m_{ij}(1_{\mathscr{R}})\alpha\,|y_j'-c|\,A(\mathscr{R}_{ij}) \leqslant V_{ij} \leqslant M_{ij}(1_{\mathscr{R}})\alpha\,|y_j''-c|\,A(\mathscr{R}_{ij})$$

donde $y_j' = y_{j-1}$ si $c \leqslant a_2$ y $y_j' = y_j$ si $c \geqslant b_2$ y $y_j'' = y_j$ si $c \leqslant a_2$ y $y_j'' = y_{j-1}$ si $c \geqslant b_2$. De aquí que si el sólido de revolución generado al hacer girar $\mathscr{R}$ un ángulo α alrededor de $y = c$ tiene volumen V_i, entonces

$$\begin{aligned} L(\alpha|I_2-c|\,1_{\mathscr{R}}, P) &= \sum_{i=1}^{k_1}\sum_{j=1}^{k_2} m_{ij}(1_{\mathscr{R}})\alpha\,|y_j'-c|\,A(\mathscr{R}_{ij}) \\ &\leqslant V \leqslant \sum_{i=1}^{k_1}\sum_{j=1}^{k_2} M_{ij}(1_{\mathscr{R}})\alpha\,|y_j''-c|\,A(\mathscr{R}_{ij}) \\ &= U(\alpha|I_2-c|\,1_{\mathscr{R}}, P). \end{aligned}$$

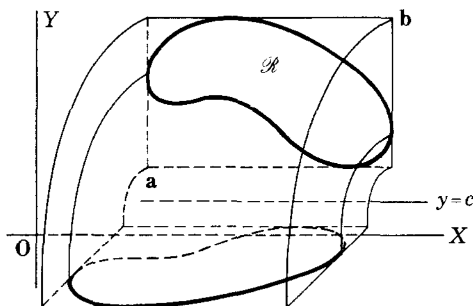


FIGURA 23

Por tanto

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} \alpha|I_2-c|\,1_{\mathscr{R}} = \underline{\int_{\mathscr{R}}} \alpha\,|y-c|\,d\mathbf{x} \leqslant V \leqslant \overline{\int_{\mathscr{R}}} \alpha\,|y-c|\,d\mathbf{x} = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} \alpha\,|I_2-c|\,1_{\mathscr{R}}.$$

Como $\mathscr{R}$ tiene área, $\alpha|I_2-c|$ es integrable sobre $\mathscr{R}$ y

12.1 $$V = \alpha\int_{\mathscr{R}} |y-c|\,d\mathbf{x}.$$

Si $\mathscr{R}$ se hace girar un ángulo α alrededor de la recta $x = c$ donde $c \leqslant a_1$ o $c \geqslant b_1$, entonces puede probarse de un modo análogo que

12.2 $$V = \alpha\int_{\mathscr{R}} |x-c|\,d\mathbf{x}.$$

Obtenemos casos especiales particularmente importantes de los anteriores resultados cuando $\alpha = 2\pi$. En los siguientes ejemplos siempre que no

hagamos mención del ángulo de revolución supondremos que $\alpha = 2\pi$ y la región ha girado una revolución completa.

12.3 Ejemplo. Encuéntrese el volumen generado cuando se hace girar alrededor del eje X la región limitada inferiormente por el eje X, en la parte superior por la curva $y = x^2$ y a la derecha por la recta $x = 2$.

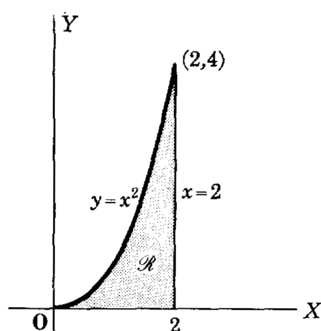


FIGURA 24

Solución. Por 12.1

$$V = 2\pi \int_{\mathscr{R}} y\, d\mathbf{x}$$

donde $\mathscr{R} = \{(x, y) \mid 0 \leqslant x \leqslant 2,\ 0 \leqslant y \leqslant x^2\}$. Por tanto,

$$\begin{aligned} V &= 2\pi \int_0^2 \int_0^{x^2} y\, dy\, dx \\ &= 2\pi \int_0^2 \left[\frac{y^2}{2}\right]_0^{x^2} dx = \pi \int_0^2 x^4\, dx \\ &= \pi \frac{x^5}{5}\bigg]_0^2 = \frac{32}{5}\pi . \end{aligned}$$

12.4 Ejemplo. Encuéntrese el volumen generado cuando se hace girar la región del ejemplo 12.3 alrededor de la recta $x = 5$.

Solución. Por 12.2, tenemos

$$\begin{aligned} V &= 2\pi \int_0^2 \int_0^{x^2} |x-5|\, dy\, dx = 2\pi \int_0^2 \int_0^{x^2} (5-x)\, dy\, dx \\ &= 2\pi \int_0^2 (5-x)x^2\, dx = 2\pi \left[\frac{5}{3}x^3 - \frac{1}{4}x^4\right]_0^2 = \frac{56\pi}{3} . \end{aligned}$$

12.5 Teorema. (Pappus, 300 A.C.) *Si a una región plana $\mathscr{R}$ la hacemos girar alrededor de una recta $\mathscr{L}$ del plano, y si $\mathscr{L}$ no intersecta a $\mathscr{R}$, entonces el volumen generado es igual al producto del área de $\mathscr{R}$ por la distancia recorrida por el centroide de $\mathscr{R}$.*

PRUEBA. Probaremos el teorema solamente para el caso especial en que $\mathscr{L}$ es una recta horizontal.

Si $\mathscr{L}$ es la recta $y = c$, entonces por 12.1

$$V = \alpha \int_{\mathscr{R}} |y-c|\, d\mathbf{x}.$$

Por otra parte (definición 10.2, pág. 356),

$$M_{\mathscr{L}} = \int_{\mathscr{R}} \operatorname{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0)\, d\mathbf{x}.$$

Como

$$\operatorname{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0) = \frac{\mathbf{n}\cdot(\mathbf{x}-\mathbf{x}_0)}{|\mathbf{n}|} = y-c,$$

tenemos

$$|M_{\mathscr{L}}| = \left|\int_{\mathscr{R}} (y-c)\, d\mathbf{x}\right| = \left|\int_{\mathscr{R}} y\, d\mathbf{x} - c\int_{\mathscr{R}} d\mathbf{x}\right|$$
$$= |M_x - cA(\mathscr{R})| = |\bar{y}A(\mathscr{R}) - cA(\mathscr{R})|.$$

Por tanto

$$V = \alpha \int_{\mathscr{R}} |y-c|\, d\mathbf{x} = \alpha \left|\int_{\mathscr{R}} (y-c)\, d\mathbf{x}\right| = \alpha |M_{\mathscr{L}}|$$
$$= \alpha |\bar{y}-c|\, A(\mathscr{R})$$

donde $\alpha|\bar{y}-c|$ es la distancia recorrida por el centroide de $\mathscr{R}$.

12.6 Ejemplo. Encuéntrese el volumen de un cono circular recto de altura h y radio r.

SOLUCIÓN. El cono puede generarse haciendo girar un triángulo rectángulo de catetos h y r alrededor del cateto de longitud h. Según el problema 5, pág. 364, el centroide está a una distancia $r/3$ del lado de longitud h. De donde, usando el teorema de Pappus, tenemos

$$V = 2\pi \frac{r}{3}\frac{rh}{2} = \frac{1}{3}\pi r^2 h$$

para el volumen del cono.

12.7 Ejemplo. Encuéntrese el volumen del toro obtenido haciendo girar un círculo de radio a alrededor de una recta $\mathscr{L}$ a una distancia b del centro del círculo ($a < b$).

SOLUCIÓN. Como el centroide de un círculo está en su centro, usando el teorema de Pappus tenemos

$$V = 2\pi b(\pi a^2) = 2\pi^2 a^2 b.$$

12.8 Ejemplo. Encuéntrese el centroide de un semicírculo.

SOLUCIÓN. Supongamos que el semicírculo es la parte superior del círculo limitado por la circunferencia $x^2+y^2 = r^2$. El eje Y es entonces un eje de simetría y, por tanto, de acuerdo a lo visto en el problema 6, pág. 364, $\bar{x} = 0$. Ahora bien, el área del semicírculo es $\frac{1}{2}\pi r^2$ y el sólido generado al hacer girar el semicírculo alrededor del eje X es una esfera de volumen $\frac{4}{3}\pi r^3$. Pero, por el teorema de Pappus

$$V = 2\pi \bar{y} A$$

luego

$$\bar{y} = \frac{V}{2\pi A} = \frac{\frac{4}{3}\pi r^3}{2\pi \frac{1}{2}\pi r^2} = \frac{4r}{3\pi}.$$

Problemas

1. Sea $\mathscr{R}$ una región plana limitada por las parábolas $y = x^2$ y $x = y^2$. Encuéntrese el volumen generado si hacemos girar $\mathscr{R}$ alrededor de:

a) el eje X *b*) la recta $x = 1$
c) la recta $y = 2$ *d*) la recta $x = -3$.

2. Una esfera de radio a tiene un hoyo cilíndrico de radio $a/2$ con un diámetro como eje. Encuéntrese el volumen del material que queda.

3. En un cilindro de 3 pulgadas de radio ha de hacerse una acanalamiento de $\frac{1}{2}$ pulgada de ancho y $\frac{1}{4}$ de pulgada de profundo. Úsese el teorema de Pappus para encontrar el volumen del material que ha de quitarse.

4. Alrededor de un cilindro de radio $3a$ se va a hacer un hendidura semicircular de radio a. Encuéntrese el volumen del material que va a quitarse.

Sugerencia. Úsese el resultado del ejemplo 12.8.

5. Un cilindro ha de recortarse en la forma que aparece en la figura 25. Encuéntrese el volumen del material que ha de quitarse.

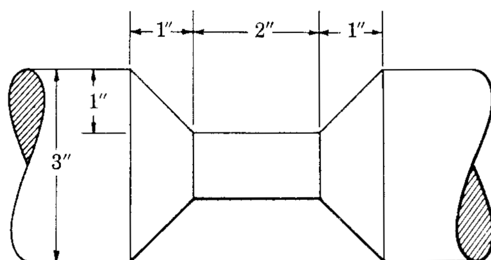


FIGURA 25

13. CAMBIO EN EL ORDEN DE INTEGRACIÓN

En algunos casos una integral iterada puede evaluarse más fácilmente si se invierte el orden de integración. Si el integrando es continuo es posible conseguir esto usando el teorema 9.3 y el corolario 9.4, pág. 355, para mostrar que la integral iterada dada es igual a una doble integral sobre alguna región y que esta doble integral es igual a una integral iterada con orden de integración diferente del orden de integración de la integral iterada original. Este procedimiento se ilustra en los siguientes ejemplos.

13.1 Ejemplo. Evalúese

$$\int_0^A \int_y^A e^{-x^2}\, dx\, dy.$$

SOLUCIÓN. No hay ninguna función elemental que tenga e^{-x^2} como su derivada, luego la integración con respecto a x no puede efectuarse en términos de funciones elementales. Si $\mathscr{R} = \{(x, y) \mid 0 \leqslant y \leqslant A,\ y \leqslant x \leqslant A\}$, entonces

$$\int_0^A \int_y^A e^{-x^2}\, dx\, dy = \int_{\mathscr{R}} e^{-x^2}\, d\mathbf{x}.$$

La región $\mathscr{R}$ es el triángulo de la figura 26.

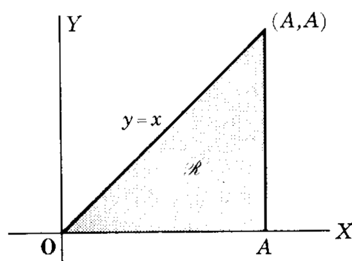


FIGURA 26

Tenemos también $\mathscr{R} = \{(x, y) \mid 0 \leqslant x \leqslant A,\ 0 \leqslant y \leqslant x\}$ de modo que

$$\int_0^A \int_y^A e^{-x^2}\,dx\,dy = \int_{\mathscr{R}} e^{-x^2}\,d\mathbf{x} = \int_0^A \int_0^x e^{-x^2}\,dy\,dx$$

$$= \int_0^A x e^{-x^2}\,dx = -\tfrac{1}{2}e^{-x^2}\Big]_0^A = \tfrac{1}{2}\left[1-e^{-A^2}\right].$$

13.2 Ejemplo. Demuéstrese que si f es continua sobre $\mathscr{R} = \{(x,y) \mid a \leqslant y \leqslant b,\ a \leqslant x \leqslant y\}$, entonces

$$\int_a^b \int_a^y f(x, y)\,dx\,dy = \int_a^b \int_x^b f(x, y)\,dy\,dx.$$

SOLUCIÓN. (Figura 27.) Si $\mathscr{R}_y = \{(x, y) \mid a \leqslant y \leqslant b,\ a \leqslant x \leqslant y\}$, entonces

$$\int_a^b \int_a^y f(x, y)\,dx\,dy = \int_{\mathscr{R}_y} f(x, y)\,d\mathbf{x}$$

y si $\mathscr{R}_x = \{(x, y) \mid a \leqslant x \leqslant b,\ x \leqslant y \leqslant b\}$, entonces

$$\int_a^b \int_x^b f(x, y)\,dy\,dx = \int_{\mathscr{R}_x} f(x, y)\,d\mathbf{x}.$$

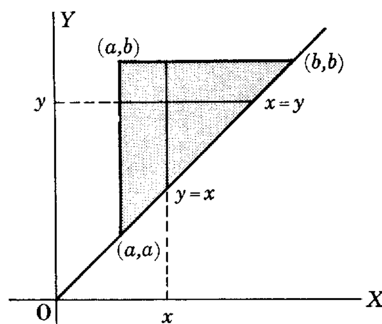


FIGURA 27

Como $\mathscr{R}_y = \mathscr{R}_x$,

$$\int_a^b \int_a^y f(x, y)\,dx\,dy = \int_a^b \int_x^b f(x, y)\,dy\,dx.$$

Problemas

1. En cada una de las siguientes integrales iteradas, inviértase el orden de integración y luego evalúese la integral iterada que parezca más sencilla.

$a)\ \int_0^1 \int_{\sqrt{y}}^1 e^{y/x}\,dx\,dy$ $\qquad b)\ \int_1^{e^2} \int_{\ln y}^2 \frac{x}{y}\,dx\,dy$

$c)\ \int_0^a \int_0^{\sqrt{a^2-x^2}} y\,dy\,dx$ $\qquad d)\ \int_0^1 \int_{\frac{1}{2}\operatorname{arccos} x}^{\pi/4} (1+x)^{-1/2}\,dy\,dx$

$e)\ \int_0^{\sqrt{3}} \int_{\arctan x}^{\pi/3} \left(1+\frac{1}{1+x^2}\right) dy\,dx$ $\quad f)\ \int_0^1 \int_{e^x}^{e} \frac{1+x}{e^x}\,dy\,dx.$

2. *a*) Úsese la relación del ejemplo 13.2 para demostrar que

$$\int_a^b \int_a^y f(x)\,dx\,dy = \int_a^b (b-x)\,f(x)\,dx.$$

b) Demuéstrese por inducción que la integral iterada *n*-ésima

$$\int_a^b \int_a^{x_n} \cdots \int_a^{x_2} f(x_1)\,dx_1 \cdots dx_{n-1}\,dx_n = \int_a^b \frac{(b-x_1)^{n-1}}{(n-1)!} f(x_1)\,dx_1\,.$$

14. INTEGRALES TRIPLES

La integral doble se introdujo en la sección 2 para intervalos en R^2 y en secciones subsecuentes el concepto se extendió a conjuntos más generales de R^2. En esta sección daremos una breve introducción a las integrales sobre conjuntos en R^3, a las que llamaremos integrales triples. El tratamiento será breve porque es análogo al tratamiento de las integrales dobles y es un caso especial de la discusión de integrales en R^n que encontraremos en la sección 19, pág. 395.

Un intervalo cerrado [**a**, **b**] en R^3 es el conjunto de todos los puntos $\mathbf{x} \in R^3$ para los cuales $a_i \leqslant x_i \leqslant b_i$ $(i = 1, 2, 3)$ donde $a_i \leqslant b_i$. Un intervalo cerrado [**a**, **b**] en R^3 es un paralelepípedo rectangular de volumen

$$V([\mathbf{a},\mathbf{b}]) = (b_1-a_1)(b_2-a_2)(b_3-a_3) = \prod_{i=1}^{3} (b_i-a_i).$$

Definimos una partición $[\mathbf{a}, \mathbf{b}] \subset R^3$ de una manera análoga a la que usamos para los intervalos en R^2. Si $\mathbf{a} = (a_1, a_2, a_3)$ y $\mathbf{b} = (b_1, b_2, b_3)$, tomamos particiones P_1 de $[a_1, b_1]$, P_2 de $[a_2, b_2]$, y P_3 de $[a_3, b_3]$. Entonces $P = P_1 \times P_2 \times P_3$ se dice que es una partición de [**a**, **b**] y la norma de P se define por

$$|P| = \text{máx}\,\{|P_1|, |P_2|, |P_3|\}.$$

Si la partición P_i subdivide $[a_i, b_i]$ en k_i subintervalos unidimensionales entonces P subdivide a [**a**, **b**] en $k = k_1 k_2 k_3$ subintervalos tridimen-

sionales. Los k subintervalos pueden enumerarse consecutivamente y denotarse por R_i ($i = 1, \ldots, k$).

Sea f una función real acotada sobre $[\mathbf{a}, \mathbf{b}]$. Como en R^2, definimos

$$m_i(f) = \text{ínf}\ \{f(\mathbf{x}) \mid \mathbf{x}\in\mathscr{R}_i\}$$
$$M_i(f) = \sup\ \{f(\mathbf{x}) \mid x\in\mathscr{R}_i\}$$

y formamos sumas inferiores y superiores:

$$L(f, P) = \sum_{i=1}^{k} m_i(f)\, V(\mathscr{R}_i)$$

$$U(f, P) = \sum_{i=1}^{k} M_i(f)\, V(\mathscr{R}_i)$$

donde $V(\mathscr{R}_i)$ es el volumen del subintervalo i-ésimo. La desigualdad

$$mV([\mathbf{a}, \mathbf{b}]) \leqslant L(f, P) \leqslant U(f, P) \leqslant MV([\mathbf{a}, \mathbf{b}]),$$

donde m y M son el ínfimo y el supremo de f sobre $[\mathbf{a}, \mathbf{b}]$ se verifica para todas las particiones P de $[\mathbf{a}, \mathbf{b}]$. Como el conjunto de todas las sumas inferiores está superiormente acotado, tiene un supremo. Análogamente, el conjunto de todas las sumas superiores tiene un ínfimo. Tenemos así integrales inferiores y superiores definidas como sigue:

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \sup\ \{L(f, P) \mid P\in\mathscr{P}\}$$

y

$$\overline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \text{ínf}\ \{U(f, P) \mid P\in\mathscr{P}\}$$

donde $\mathscr{P}$ es el conjunto de todas las particiones de $[\mathbf{a}, \mathbf{b}]$. De nuevo nuestro maximo interés recae en aquellas funciones para las que las integrales inferior y superior coinciden. Al valor común se le llama la *integral (de Riemann) de f sobre* $[\mathbf{a}, \mathbf{b}]$, y se denota por $\int_{\mathbf{a}}^{\mathbf{b}} f$ o $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\, d\mathbf{x}$. La integral de una función f sobre un intervalo $[\mathbf{a}, \mathbf{b}]$ en R^3 se llama una integral triple. El término integral triple se deriva de la naturaleza tridimensional del intervalo de integración. Las notaciones

$$\iiint_{\mathbf{a}}^{\mathbf{b}} f(x, y, z)\, dx\, dy\, dz \quad \text{e} \quad \iiint_{\mathbf{a}}^{\mathbf{b}} f(x, y, z)\, dV$$

son a veces usadas para denotar a la integral triple de f sobre $[\mathbf{a}, \mathbf{b}]$.

Se puede mostrar que la integral triple tiene las siguientes propiedades básicas:

14.1 *Si f es integrable sobre $[\mathbf{a}, \mathbf{b}]$ y c es una función constante de* R^3 *en* R, *entonces la función cf es integrable e* $\int_{\mathbf{a}}^{\mathbf{b}} cf = c \int_{\mathbf{a}}^{\mathbf{b}} f$.

14.2 *Si c es una función constante de* R^3 *en* R, *entonces para todo intervalo* $[\mathbf{a}, \mathbf{b}] \in R^3$, *$c$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$ e* $\int_{\mathbf{a}}^{\mathbf{b}} c = cV([\mathbf{a}, \mathbf{b}])$.

14.3 *Si las funciones f y g son integrables sobre $[\mathbf{a}, \mathbf{b}]$ entonces $f+g$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$ e* $\int_{\mathbf{a}}^{\mathbf{b}} (f+g) = \int_{\mathbf{a}}^{\mathbf{b}} f + \int_{\mathbf{a}}^{\mathbf{b}} g$.

14.4 *La función f es integrable sobre $[\mathbf{a}, \mathbf{b}]$ si y sólo si las funciones f^+ y f^- definidas por las reglas*

$$f^+(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & si\ f(\mathbf{x}) \geqslant 0 \\ 0 & si\ f(\mathbf{x}) < 0 \end{cases} \quad y \quad f^-(\mathbf{x}) = \begin{cases} 0 & si\ f(\mathbf{x}) \geqslant 0 \\ -f(\mathbf{x}) & si\ f(\mathbf{x}) < 0 \end{cases}$$

son, ambas, integrables sobre $[\mathbf{a}, \mathbf{b}]$.

14.5 *Si la función f es integrable sobre $[\mathbf{a}, \mathbf{b}]$, entonces f^2 es integrable sobre $[\mathbf{a}, \mathbf{b}]$.*

14.6 *Si las funciones f y g son integrables sobre $[\mathbf{a}, \mathbf{b}]$, entonces el producto fg es integrable sobre $[\mathbf{a}, \mathbf{b}]$.*

14.7 *Si las funciones f y g son integrables sobre $[\mathbf{a}, \mathbf{b}]$ y $f(\mathbf{x}) \leqslant g(\mathbf{x})$ para toda $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$, entonces* $\int_{\mathbf{a}}^{\mathbf{b}} f \leqslant \int_{\mathbf{a}}^{\mathbf{b}} g$.

14.8 *Si la función f es integrable sobre $[\mathbf{a}, \mathbf{b}]$, entonces $|f|$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$ e* $\left| \int_{\mathbf{a}}^{\mathbf{b}} f \right| \leqslant \int_{\mathbf{a}}^{\mathbf{b}} |f|$.

La integral de una función f sobre un conjunto acotado $\mathscr{E} \subset R^3$ está definida en términos de la función $f_{\mathscr{E}}$ donde

$$f_{\mathscr{E}}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{para} \quad \mathbf{x} \in \mathscr{E} \\ 0 & \text{para} \quad \mathbf{x} \in \mathscr{C}\mathscr{E} \end{cases}$$

donde $\mathscr{C}\mathscr{E}$, el complemento de $\mathscr{E}$, es el conjunto de todos los puntos de R^3 no en $\mathscr{E}$. Si $[\mathbf{a}, \mathbf{b}]$ es un intervalo en R^3 tal que $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$ y si $\int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}}$ existe, entonces $\int_{\mathscr{E}} f = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}}$. La integral no depende del intervalo $[\mathbf{a}, \mathbf{b}]$ escogido

para contener $\mathscr{E}$. Las propiedades básicas 14.1-14.8 puede demostrarse que son válidas para $\int_{\mathscr{E}} f$.

Si $\mathscr{E}$ es un conjunto acotado en R^3 y $[\mathbf{a}, \mathbf{b}]$ es un intervalo en R^3 tal que $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$, entonces $\underline{V}(\mathscr{E}) = \int_{\mathbf{a}}^{\mathbf{b}} 1_{\mathscr{E}_i}$ y $\overline{V}(\mathscr{E}) = \int_{\mathbf{a}}^{\mathbf{b}} 1_{\overline{\mathscr{E}}}$ se llaman, respectivamente, *volumen interior* y *volumen exterior de* $\mathscr{E}$. Si $\underline{V}(\mathscr{E}) = \overline{V}(\mathscr{E})$ entonces $1_{\mathscr{E}}$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$ y el *volumen de* $\mathscr{E}$ es, por definición, este valor común y se sigue que

$$V(\mathscr{E}) = \int_{\mathbf{a}}^{\mathbf{b}} 1_{\mathscr{E}} = \int_{\mathscr{E}} 1 .$$

Si un conjunto $\mathscr{E}$ tiene volumen, entonces la frontera de $\mathscr{E}$ tiene volumen cero (véase el teorema 4.17, pág. 337).

Fácilmente se puede probar (problema 3) que el volumen tiene propiedades análogas a las propiedades del área dadas en los teoremas 4.8, 4.10, y corolario 4.13 (pág. 333).

Si $\mathscr{E}$ *y* $\mathscr{F}$ *son conjuntos en* R^3 *que tienen volumen, entonces*

14.9 $V(\mathscr{E}) \geqslant 0$.

14.10 *Si* $\mathscr{E} \subset \mathscr{F}$, *entonces* $V(\mathscr{E}) \leqslant V(\mathscr{F})$.

14.11 *Si* $V(\mathscr{E} \cap \mathscr{F}) = 0$, *entonces* $V(\mathscr{E} \cup \mathscr{F}) = V(\mathscr{E}) + V(\mathscr{F})$.

14.12 Teorema. *Si f es una función acotada sobre un intervalo* $[\mathbf{a}, \mathbf{b}] \subset R^3$ *y el conjunto* $\mathscr{E}$ *de puntos de discontinuidad de f sobre* $[\mathbf{a}, \mathbf{b}]$ *tiene volumen cero, entonces* $\int_{\mathbf{a}}^{\mathbf{b}} f$ *existe.*

Como este teorema es análogo al teorema 5.1 (pág. 338), omitimos aquí su prueba.

Del teorema 14.12 se siguen varios importantes corolarios. Estos corolarios son análogos a los corolarios de la sección 5.

14.13 Corolario. *Sean f y g dos funciones acotadas sobre un intervalo* $[\mathbf{a}, \mathbf{b}] \subset R^3$ *y continuas sobre* $[\mathbf{a}, \mathbf{b}] - \mathscr{E}$, *donde* $\mathscr{E}$ *es un subconjunto de* $[\mathbf{a}, \mathbf{b}]$ *de volumen cero. Si $f = g$ sobre* $[\mathbf{a}, \mathbf{b}] - \mathscr{E}$, *entonces*

$$\int_{\mathbf{a}}^{\mathbf{b}} f = \int_{\mathbf{a}}^{\mathbf{b}} g .$$

14.14 Corolario. *Un conjunto acotado* $\mathscr{E} \subset R^3$ *tiene volumen si la frontera de* $\mathscr{E}$ *tiene volumen cero.*

14.15 Corolario. *Sea $\mathscr{E}$ un volumen acotado en* R^3 *que tiene volumen. Si f está acotada sobre $\mathscr{E}$ y es continua en el interior de $\mathscr{E}$, entonces* $\int_{\mathscr{E}} f$ *existe.*

14.16 Corolario. *Sea $\mathscr{E}$ un conjunto acotado en* R^3 *que tiene volumen. Si f está acotada sobre $\mathscr{E}$ y es continua en el interior de $\mathscr{E}$, excepto sobre un subconjunto $\mathscr{F}$ de volumen cero, entonces* $\int_{\mathscr{E}} f$ *existe.*

14.17 Corolario. *Sea $\mathscr{E}$ un conjunto cerrado y acotado en* R^3 *que tiene volumen. Si f es continua sobre $\mathscr{E}$, entonces* $\int_{\mathscr{E}} f$ *existe.*

Problemas

1. Pruébese que los siguientes conjuntos tienen volumen cero.
   a) un número finito de puntos
   b) un segmento rectilíneo
   c) un subconjunto acotado de un plano.
2. Pruébese que una esfera tiene volumen.
3. Demuéstrese que el volumen tiene las propiedades 14.9, 14.10 y 14.11.

15. INTEGRALES ITERADAS

Las integrales iteradas en dos dimensiones se consideraron en la sección 7. Aquí consideraremos integrales iteradas en tres dimensiones.

La integral iterada

15.1
$$\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,y)}^{h_2(x,y)} f(x,y,z)\,dz\,dy\,dx$$

se interpreta como $\int_a^b F(x)\,dx$ donde para cada $x \in [a,b]$

$$F(x) = \int_{g_1(x)}^{g_2(x)} G(x,y)\,dy$$

y para cada $(x,y) \in \{(x,y) \mid a \leqslant x \leqslant b,\ g_1(x) \leqslant y \leqslant g_2(x)\}$

$$G(x,y) = \int_{h_1(x,y)}^{h_2(x,y)} f(x,y,z)\,dz\,.$$

Es decir

$$\int_a^b F(x)\,dx = \int_a^b \left\{\int_{g_1(x)}^{g_2(x)} G(x,y)\,dy\right\}dx$$

$$= \int_a^b \left\{\int_{g_1(x)}^{g_2(x)} \left\{\int_{h_1(x,y)}^{h_2(x,y)} f(x,y,z)\,dz\right\}dy\right\}dx.$$

15.2 Ejemplo. Evalúese

$$\int_0^a \int_0^{b(1-x/a)} \int_0^{c(1-x/a-y/b)} x\,dz\,dy\,dx.$$

SOLUCIÓN.

$$\int_0^a \int_0^{b(1-x/a)} \int_0^{c(1-x/a-y/b)} x\,dz\,dy\,dx = \int_0^a \int_0^{b(1-x/a)} [xz]_0^{c(1-x/a-y/b)}\,dy\,dx$$

$$= \int_0^a \int_0^{b(1-x/a)} c\left[x\left(1-\frac{x}{a}\right)-\frac{xy}{b}\right]dy\,dx$$

$$= c\int_0^a \left[x\left(1-\frac{x}{a}\right)y-\frac{xy^2}{2b}\right]_0^{b(1-x/a)} dx$$

$$= c\int_0^a \left[bx\left(1-\frac{x}{a}\right)^2-\tfrac{1}{2}bx\left(1-\frac{x}{a}\right)^2\right]dx$$

$$= \tfrac{1}{2}bc\int_0^a x\left(1-\frac{x}{a}\right)^2 dx$$

$$= \tfrac{1}{2}bc\left[\frac{x^2}{2}-\frac{2x^3}{3a}+\frac{x^4}{4a^2}\right]_0^a = \frac{a^2bc}{24}.$$

Problemas

Evalúense las siguientes integrales iteradas.

a) $\int_0^5 \int_0^2 \int_1^3 (x^2+2yz)\,dx\,dy\,dz$

b) $\int_0^1 \int_{\sqrt{x}}^{x^2} \int_{y^2}^{x^2} z\,dz\,dy\,dx$

c) $\int_{-1}^2 \int_0^{\pi/2} \int_0^{y^2} \frac{z}{y}\,\text{sen}\,\frac{x}{y}\,dx\,dy\,dz$

d) $\displaystyle\int_0^1 \int_0^{\sqrt{1-x^2}} \int_0^{(1-x^2-y^2)^{3/4}} (1-x^2-y^2)^{1/4}\, dz\, dy\, dx$

e) $\displaystyle\int_0^a \int_0^{b(1-x/a)} \int_0^{c(1-x/a-y/b)} dz\, dy\, dx$

f) $\displaystyle\int_0^a \int_0^{b\sqrt{1-x^2/a^2}} \int_0^{c\sqrt{1-x^2/a^2-y^2/b^2}} dz\, dy\, dx$

g) $\displaystyle\int_0^a \int_0^{b(1-x/a)} \int_0^{c(1-x/a-y/b)} z\, dz\, dy\, dx$

h) $\displaystyle\int_0^{2\pi} \int_0^a \int_0^{\sqrt{a^2-r^2}} r\, dz\, dr\, d\theta .$

16. TEOREMA FUNDAMENTAL PARA LAS INTEGRALES TRIPLES

En esta sección enunciaremos un teorema que establece una relación fundamental entre las integrales triples y las integrales iteradas. Este teorema nos muestra un importante método para la evaluación de las integrales triples. Este teorema es el análogo para los intervalos tridimensionales del teorema 8.1, pág. 347, para los intervalos bidimensionales.

16.1 Teorema. *Si* $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ *existe y* $\mathbf{a} = (a_1, a_2, a_3)$ *y* $\mathbf{b} = (b_1, b_2, b_3)$ *y si* $F(x, y) = \int_{a_3}^{b_3} f(x, y, z)\,dz$ *existe para todo* $(x, y) \in [\mathbf{c}, \mathbf{d}]$ *con* $\mathbf{c} = (a_1, a_2)$ *y* $\mathbf{d} = (b_1, b_2)$, *entonces*

$$\int_{\mathbf{c}}^{\mathbf{d}} \int_{a_3}^{b_3} f(x, y, z)\,dz\,dA = \int_{\mathbf{c}}^{\mathbf{d}} F(x, y)\,dA$$

existe e

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{\mathbf{c}}^{\mathbf{d}} \int_{a_3}^{b_3} f(x, y, z)\,dz\,dA .$$

La prueba del teorema 16.1 es esencialmente la misma que la prueba del teorema 8.1 y no la daremos.

El siguiente corolario da condiciones suficientes para asegurar la existencia de las integrales del teorema 16.1.

16.2 Corolario. *Sea* f *una función acotada sobre el intervalo* $[\mathbf{a}, \mathbf{b}]$ *en* R^3 *y continua sobre* $[\mathbf{a}, \mathbf{b}]$ *excepto sobre un conjunto* $\mathscr{E}$ *de volumen cero con la propiedad de que cada recta paralela al eje* Z *intersecta a* $\mathscr{E}$ *en cuando más*

un número finito de puntos. Entonces las integrales del teorema 16.1 existen e

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{\mathbf{c}}^{\mathbf{d}} \int_{a_3}^{b_3} f(x, y, z)\,dz\,dA\,.$$

La prueba de este corolario es esencialmente la misma que la prueba del teorema 8.8, pág. 351.

16.3 Corolario. *Si* i) $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ *existe,* ii) $F(x, y) = \int_{a_3}^{b_3} f(x, y, z)\,dz$ *existe para todo* $(x, y) \in [\mathbf{c}, \mathbf{d}]$ *donde* $\mathbf{c} = (a_1, a_2)$ *y* $\mathbf{d} = (b_1, b_2)$, *e* iii) $G(x) = \int_{a_2}^{b_2} F(x, y)\,dy$ *existe para todo* $x \in [a_1, b_1]$, *entonces* $\int_{a_1}^{b_1} G(x)\,dx$ *existe e*

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{a_1}^{b_1} G(x)\,dx = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} f(x, y, z)\,dz\,dy\,dx\,.$$

PRUEBA. Por el teorema 16.1, la existencia de $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ y la existencia de $F(x, y) = \int_{a_3}^{b_3} f(x, y, z)\,dz$ para todo $(x, y) \in [\mathbf{c}, \mathbf{d}]$ implica la existencia de $\int_{\mathbf{c}}^{\mathbf{d}} F(x, y)\,dA$ e

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{\mathbf{c}}^{\mathbf{d}} F(x, y)\,dA = \int_{\mathbf{c}}^{\mathbf{d}} \int_{a_3}^{b_3} f(x, y, z)\,dz\,dA\,.$$

Ahora bien, según el teorema 8.1, la existencia de $\int_{\mathbf{c}}^{\mathbf{d}} F(x, y)\,dA$ y de $G(x) = \int_{a_2}^{b_2} F(x, y)\,dy$ para toda $x \in [a_1, b_1]$ implica que $\int_{a_1}^{b_1} G(x)\,dx$ existe e

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{\mathbf{c}}^{\mathbf{d}} F(x, y)\,dA = \int_{a_1}^{b_1} G(x)\,dx$$

$$= \int_{a_1}^{b_1} \int_{a_2}^{b_2} F(x, y)\,dy\,dx = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} f(x, y, z)\,dz\,dy\,dx\,.$$

16.4 Corolario. *Si f es continua sobre* $[\mathbf{a}, \mathbf{b}]$, *entonces*

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x} = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} f(x, y, z)\,dz\,dy\,dx\,.$$

PRUEBA. La continuidad de f implica la existencia y continuidad de F y G del corolario 16.3 y de ello sigue el resultado.

Enunciamos ahora, sin prueba, un teorema para regiones $\mathscr{R}_{xy}$ que es análogo al corolario 16.4 para intervalos.

16.5 Teorema. *Si f es continua sobre la región $\mathscr{R}_{xy} = \{(x, y, z) \mid a \leqslant x \leqslant b, g_1(x) \leqslant y \leqslant g_2(x), h_1(x, y) \leqslant z \leqslant h_2(x, y)\}$ donde g_1, g_2 son continuas sobre $[a, b]$ y h_1, h_2 son continuas sobre $\mathscr{R}_x = \{(x, y) \mid a \leqslant x \leqslant b, g_1(x) \leqslant y \leqslant g_2(x)\}$, entonces*

$$\int_{\mathscr{R}_{xy}} f(\mathbf{x})\, d\mathbf{x} = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,y)}^{h_2(x,y)} f(x, y, z)\, dz\, dy\, dx .$$

Resultados análogos a los del teorema 16.5 pueden obtenerse para las varias permutaciones de x, y, z.

Nota. En la evaluación de una integral triple sobre $\mathscr{R}_{xy}$ por medio de la integral iterada del teorema 16.5, nótese que integramos primero con respecto a z desde la superficie límite inferior a la superficie límite superior y luego con respecto a x y y sobre la proyección de $\mathscr{R}_{xy}$ sobre el plano XY.

16.6 Ejemplo. Evalúese la integral

$$\int_{(0,0,0)}^{(2,5,3)} (x^2 + y^2 + z^2)\, d\mathbf{x} .$$

SOLUCIÓN. Como $f = I_1{}^2 + I_2{}^2 + I_3{}^2$ es continua sobre $[(0, 0, 0), (2, 5, 3)]$, según el corolario 16.4 tenemos

$$\int_{(0,0,0)}^{(2,5,3)} (x^2 + y^2 + z^2)\, d\mathbf{x} = \int_0^2 \int_0^5 \int_0^3 (x^2 + y^2 + z^2)\, dz\, dy\, dx = 380.$$

16.7 Ejemplo. Evalúese la integral $\int_{\mathscr{R}} x^2\, d\mathbf{x}$ donde $\mathscr{R}$ es el tetraedro limitado por los planos coordenados y el plano $2x + 3y + z = 6$.

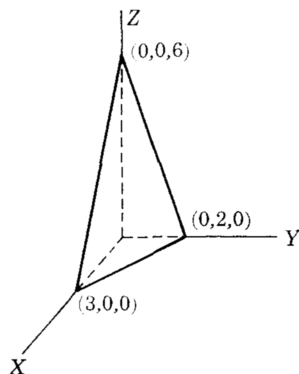


FIGURA 28

SOLUCIÓN. (Figura 28.) $\mathscr{R}$ está limitada hacia arriba por $z = 6-2x-3y$ y hacia abajo por el plano $z = 0$. La proyección $\mathscr{R}_x$ de $\mathscr{R}$ paralela al eje Z es el triángulo limitado por la recta $2x+3y = 6$ y los ejes coordenados: $x = 0$, $y = 0$. De donde

$$\int_{\mathscr{R}} x^2\,d\mathbf{x} = \int_{\mathscr{R}_x}\left[\int_0^{6-2x-3y} x^2\,dz\right]dA$$

$$= \int_0^3 \int_0^{\frac{1}{3}(6-2x)} \int_0^{6-2x-3y} x^2\,dz\,dy\,dx = \tfrac{27}{5}.$$

Problemas

1. Encuéntrese $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\,d\mathbf{x}$ para cada una de las siguientes funciones.

a) $f(\mathbf{x}) = xyz$, $\mathbf{a} = (0, 2, 4)$, $\mathbf{b} = (1, 6, 5)$
b) $f(\mathbf{x}) = x^2$, $\mathbf{a} = (-1, 2, -3)$, $\mathbf{b} = (1, 3, -2)$
c) $f(\mathbf{x}) = xz\cos(xy)$, $\mathbf{a} = (0, 2, 1)$, $\mathbf{b} = (2, 5, 2)$
d) $f(\mathbf{x}) = \operatorname{sen}(x+y)$, $\mathbf{a} = (0, 1, 1)$, $\mathbf{b} = (2, 2, 2)$
e) $f(x) = |\mathbf{x}|^2$, $\mathbf{a} = (0, 0, 0)$, $\mathbf{b} = (1, 1, 1)$
f) $f(\mathbf{x}) = z^2 \operatorname{sen} y$, $\mathbf{a} = (0, 0, 0)$, $\mathbf{b} = (2\pi, \pi, a)$.

2. Evalúense las siguientes integrales:

a) $\int_{\mathscr{R}} d\mathbf{x}$ con $\mathscr{R} = \{(x, y, z)\,|\,0 \leqslant x \leqslant a,\ 0 \leqslant y \leqslant \sqrt{a^2-x^2},\ 0 \leqslant z \leqslant \sqrt{a^2-x^2-y^2}\}$

b) $\int_{\mathscr{R}} x\,d\mathbf{x}$ con $\mathscr{R} = \left\{(x, y, z)\,|\,0 \leqslant x \leqslant a,\ 0 \leqslant y \leqslant b\left(1-\frac{x}{a}\right),\ 0 \leqslant z \leqslant c\left(1-\frac{x}{a}-\frac{y}{b}\right)\right\}$

c) $\int_{\mathscr{R}} \frac{y}{y^2+z^2}\,d\mathbf{x}$ con $\mathscr{R} = \{(x, y, z)\,|\,2 \leqslant x \leqslant 3,\quad 0 \leqslant y \leqslant x,\quad 0 \leqslant z \leqslant y\}$

d) $\int_{\mathscr{R}} yz\,d\mathbf{x}$ con $\mathscr{R} = \{(x, y, z)\,|\,0 \leqslant x \leqslant 4,\ 0 \leqslant y \leqslant \sqrt{x},\ 0 \leqslant z \leqslant y\}$

e) $\int_{\mathscr{R}} x^2 y\,d\mathbf{x}$ con $\mathscr{R} = \left\{(x, y, z)\,|\,0 \leqslant z \leqslant \frac{\pi}{2},\ 0 \leqslant y \leqslant a,\ 0 \leqslant x \leqslant y\cos z\right\}$

f) $\int_{\mathscr{R}} d\mathbf{x}$ con $\mathscr{R} = \{(x, y, z) \mid 0 \leqslant x \leqslant 2, 0 \leqslant y \leqslant \sqrt{2-\frac{1}{2}x^2},$
$x^2+3y^2 \leqslant z \leqslant 8-x^2-y^2\}$

3. Evalúese $\int_{\mathscr{R}} x\, d\mathbf{x}$ donde $\mathscr{R}$ está superiormente acotado por $z = 16-4x^2-y^2$ e inferiormente acotada por $z = 12x^2+y^2$.

4. Evalúese $\int_{\mathscr{R}} d\mathbf{x}$ donde $\mathscr{R}$ es la región en el primer octante ($x \geqslant 0$, $y \geqslant 0$, $z \geqslant 0$) limitada por el elipsoide $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$.

17. APLICACIONES DE LAS INTEGRALES TRIPLES

En la sección 14 indicamos que el volumen de un conjunto $\mathscr{E}$ en R^3 es $\int_{\mathscr{E}} 1$ para aquellos conjuntos que tienen volumen. También se señaló que si $\mathscr{E}$ y $\mathscr{F}$ son conjuntos en R^3 que tienen volumen, entonces $V(\mathscr{E}) \geqslant 0$; si $\mathscr{E} \subset \mathscr{F}$, entonces $V(\mathscr{E}) \leqslant V(\mathscr{F})$; y si $V(\mathscr{E} \cap \mathscr{F}) = 0$, entonces $V(\mathscr{E} \cup \mathscr{F}) = V(\mathscr{E}) + V(\mathscr{F})$. Antes, en la sección 11, se probó que si el volumen tiene estas propiedades, entonces, para regiones $\mathscr{R}_{xy}$, en R^3 del tipo

$$\mathscr{R}_{xy} = \{(x, y, z) \mid 0 \leqslant z \leqslant h(x, y), (x, y) \in \mathscr{R}_x\}$$

donde $\mathscr{R}_x = \{(x, y) \mid a \leqslant x \leqslant b, g_1(x) \leqslant y \leqslant g_2(x)\}$, tenemos

$$V(\mathscr{R}_{xy}) = \int_{\mathscr{R}_x} h(x, y)\, dx\, dy .$$

Por el teorema 16.5,

$$\int_{\mathscr{R}_{xy}} 1 = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_0^{h(x,y)} dz\, dy\, dx = \int_a^b \int_{g_1(x)}^{g_2(x)} h(x, y)\, dy\, dx$$
$$= \int_{\mathscr{R}_x} h(x, y)\, dx\, dy .$$

Así pues, para regiones del tipo $\mathscr{R}_{xy}$ estos dos métodos de determinación del volumen coinciden y son esencialmente el mismo.

17.1 Ejemplo. Encuéntrese el volumen del tetraedro limitado por el plano $\dfrac{x}{a} + \dfrac{y}{b} + \dfrac{z}{c} = 1$ (a, b, c todos positivos) y los planos coordenados.

SOLUCIÓN. (Figura 29.) Aquí la región es

$$\mathscr{R}_{xy} = \left\{(x, y, z) \mid (x, y) \in \mathscr{R}_x,\ 0 \leqslant z \leqslant c\left(1 - \frac{x}{a} - \frac{y}{b}\right)\right\}$$

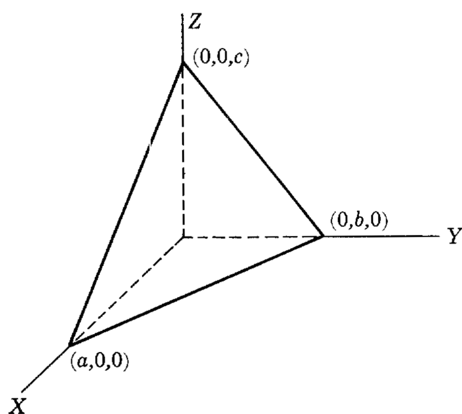


FIGURA 29

donde

$$\mathscr{R}_x = \left\{(x, y) \mid 0 \leqslant x \leqslant a,\ 0 \leqslant y \leqslant b\left(1 - \frac{x}{a}\right)\right\}.$$

De donde

$$\text{volumen} = \int_{\mathscr{R}_{xy}} 1 = \int_{\mathscr{R}_x} \left\{\int_0^{c(1-x/a-y/b)} dz\right\} dA$$

$$= \int_0^a \int_0^{b(1-x/a)} \int_0^{c(1-x/a-y/b)} dz\, dy\, dx$$

$$= \int_0^a \int_0^{b(1-x/a)} c\left(1 - \frac{x}{a} - \frac{y}{b}\right) dy\, dx$$

$$= \frac{bc}{2} \int_0^a \left(1 - \frac{x}{a}\right)^2 dx = \frac{abc}{6}.$$

Las definiciones de primer momento y segundo momento de regiones en R^3 con respecto a planos son análogas a las definiciones dadas en la definición 10.2, pág. 356, para momentos de regiones planas con respecto a rectas.

17.2 Definición. *Sea $\mathscr{P}$ un plano en* R^3 *con normal* $\mathbf{n}$ *y sea* $\mathbf{x}_0$ *un punto fijo de $\mathscr{P}$. El (primer)* ***momento*** $M_{\mathscr{P}}$ *y el* ***momento de inercia*** $I_{\mathscr{P}}$ *de una región* R *en* R^3 *con respecto al plano $\mathscr{P}$ están definidos por*

$$M_{\mathscr{P}} = \int_{\mathscr{R}} \text{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0)\, d\mathbf{x}$$

e

$$I_{\mathscr{P}} = \int_{\mathscr{R}} [\text{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0)]^2\, d\mathbf{x}$$

respectivamente.

En la definición de $M_{\mathscr{P}}$ e $I_{\mathscr{P}}$ cuando $\mathscr{P}$ es el plano XY, es habitual escoger $\mathbf{n}$ en la dirección positiva del eje Z. Entonces

$$\text{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0) = \mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_0) = z$$

y

$$M_{xy} = \int_{\mathscr{R}} z\, d\mathbf{x}, \qquad I_{xy} = \int_{\mathscr{R}} z^2\, d\mathbf{x}.$$

Cuando $\mathscr{P}$ es el plano YZ, se escoge $\mathbf{n}$ en la dirección positiva del eje X. Entonces

$$\text{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0) = \mathbf{i}\cdot(\mathbf{x}-\mathbf{x}_0) = x$$

y

$$M_{yz} = \int_{\mathscr{R}} x\, d\mathbf{x}, \qquad I_{yz} = \int_{\mathscr{R}} x^2\, d\mathbf{x}.$$

Cuando $\mathscr{P}$ es el plano XZ, escogemos $\mathbf{n}$ en la dirección positiva del eje Y. Entonces

$$\text{Comp}_{\mathbf{n}}(\mathbf{x}-\mathbf{x}_0) = \mathbf{j}\cdot(\mathbf{x}-\mathbf{x}_0) = y$$

y

$$M_{zx} = \int_{\mathscr{R}} y\, d\mathbf{x}, \qquad I_{zx} = \int_{\mathscr{R}} y^2\, d\mathbf{x}.$$

El momento de inercia de una región $\mathscr{R}$ en R^3 con respecto a una recta $\mathscr{L}$ puede definirse como la suma de los momentos de inercia con respecto a cualquier par de planos ortogonales que tengan a $\mathscr{L}$ como su recta de intersección.

17.3 Definición. *El punto*

$$\bar{\mathbf{x}} = (\bar{x}, \bar{y}, \bar{z}) = \left(\frac{M_{yz}}{V(\mathscr{R})}, \frac{M_{zx}}{V(\mathscr{R})}, \frac{M_{xy}}{V(\mathscr{R})}\right)$$

se llama ***centroide*** *de $\mathscr{R}$.*

17.4 Ejemplo. Encuéntrense los momentos primero y segundo con respecto a los planos coordenados del tetraedro limitado por el plano $\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1$ y los planos de coordenadas. Encuéntrense también el centroide y los segundos momentos con respecto a los ejes de coordenadas.

SOLUCIÓN. (Figura 29.) Esta es la región del ejemplo 17.1:

$$\mathscr{R}_{xy}=\left\{(x,y,z)\,|\,0\leqslant x\leqslant a,\ 0\leqslant y\leqslant b\left(1-\frac{x}{a}\right),\ 0\leqslant z\leqslant c\left(1-\frac{x}{a}-\frac{y}{b}\right)\right\}.$$

De donde

$$M_{yz}=\int_{\mathscr{R}_{xy}} x\,d\mathbf{x}=\int_0^a\int_0^{b(1-x/a)}\int_0^{c(1-x/a-y/b)} x\,dz\,dy\,dx\,.$$

Las integraciones con respecto a z y y son exactamente las mismas que las dadas para el volumen en el ejemplo 17.1 excepto en que el integrando aparece multiplicado por x.

$$M_{yz}=\frac{bc}{2}\int_0^a x\left(1-\frac{x}{a}\right)^2 dx=\frac{bc}{2a^2}\int_0^a (a^2x-2ax^2+x^3)\,dx$$

$$=\frac{bc}{2a^2}\left[\frac{a^2x^2}{2}-\frac{2ax^3}{3}+\frac{x^4}{4}\right]_0^a=\frac{a^2bc}{24}\,.$$

Para encontrar M_{zx} observamos que no es necesario efectuar de nuevo la integración. Si reemplazamos (x,y,z) por (y,z,x) y (a, b, c) por (b, c, a) en el cálculo de M_{yz}, obtenemos

$$M_{zx}=\frac{ab^2c}{24}\,.$$

Análogamente,

$$M_{xy}=\frac{abc^2}{24}\,.$$

Para los segundos momentos tenemos,

$$I_{yz}=\int_0^a\int_0^{b(1-x/a)}\int_0^{c(1-x/a-y/b)} x^2\,dz\,dy\,dx$$

$$=\frac{bc}{2a^2}\int_0^a (a^2x^2-2ax^3+x^4)\,dx=\frac{a^3bc}{60}\,.$$

De nuevo, como en el caso de los primeros momentos, por intercambio cíclico de (x, y, z) y (a, b, c) obtenemos,

$$I_{zx} = \frac{ab^3c}{60} \quad \text{e} \quad I_{xy} = \frac{abc^3}{60}.$$

Como $V(\mathscr{R}) = \frac{1}{6}abc$, tenemos

$$\bar{\mathbf{x}} = \left(\frac{a}{4}, \frac{b}{4}, \frac{c}{4}\right).$$

Los segundos momentos con respecto a los ejes coordenados son

$$I_x = \int_{\mathscr{R}_{xy}} (y^2+z^2)\,d\mathbf{x} = \int_{\mathscr{R}_{xy}} y^2\,d\mathbf{x} + \int_{\mathscr{R}_{xy}} z^2\,d\mathbf{x} = I_{zx}+I_{xy}$$

$$= \frac{ab^3c}{60} + \frac{abc^3}{60} = \frac{abc}{60}(b^2+c^2),$$

$$I_y = \frac{abc}{60}(a^2+c^2),$$

e

$$I_z = \frac{abc}{60}(a^2+b^2).$$

Problemas

1. Encuéntrese el volumen del elipsoide

2. Encuéntrese el centroide de la porción del elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ que se encuentra en el primer octante.

3. Encuéntrense los momentos de inercia con respecto a cada uno de los planos coordenados de la región del problema 2.

4. Encuéntrese el volumen de la región limitada por los paraboloides $z = 4x^2+y^2$ y $z = 4-x^2-4y^2$

5. Encuéntrese el volumen de la región limitada por las superficies $x^2+y^2 = 4$, $z = 4-y^2$ y $z = 0$.

6. Encuéntrese el centroide de la región del problema 5.

7. Encuéntrese el volumen de la región situada en el primer octante limitada por los cilindros $x^2+z^2 = a^2$ y $y^2+z^2 = a^2$.

8. Encuéntrese el centroide de la región del problema 7.

*9. Encuéntrese el volumen de la región limitada por el cono elíptico $\frac{x^2}{a^2}+\frac{y^2}{b^2} = \frac{z^2}{c^2}$ y el plano $z = c$.

18. ÁREA, VOLUMEN Y MOMENTOS SIN INTEGRACIÓN

En muchos casos es posible encontrar los primeros y los segundos momentos, los centroides, las áreas y los volúmenes de ciertas regiones sin necesidad de recurrir a la integración. Si se conocen el área y los momentos para los rectángulos, los triángulos y los semicírculos (y cuartos de círculos) entonces el área, los momentos y los centroides de las regiones que consistan en combinaciones de tales figuras pueden encontrarse sin dificultad. Además, por el uso del teorema de Pappus, pueden calcularse fácilmente también los volúmenes de revolución que se obtienen por el giro de tales regiones. A continuación enumeramos algunos de los resultados que ya previamente se habían obtenido respecto al rectángulo, al círculo y al semicírculo en la tabla I. En esta tabla, CG denota al centroide (CG, por centro de gravedad) y J_p denota al momento polar de inercia con respecto al centroide. I_{AA}, etc., representan los momentos de inercia respecto a los ejes AA, etc. que aparecen marcados en las correspondientes figuras.

18.1 Ejemplo. Encuéntrense el centroide y el momento de inercia con respecto a la base de la región en forma de L que se muestra en la figura 30.

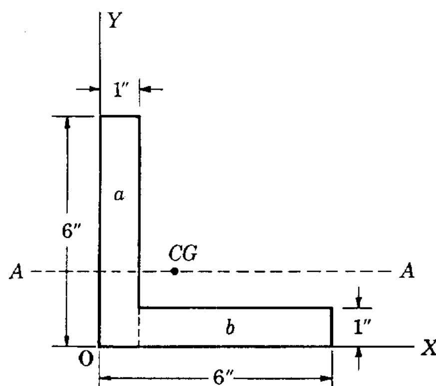


FIGURA 30

Tabla I

Región	Centroide	Momento de inercia
Rectángulo **Rectangle**	$\bar{y} = \frac{h}{2}$ $\bar{x} = \frac{b}{2}$	$I_{AA} = \frac{bh^3}{12}$ $I_{BB} = \frac{b^3h}{12}$ $I_{CC} = \frac{bh^3}{3}$ $J_p = \frac{bh^3 + b^3h}{12}$
Triángulo **Triangle**	$\bar{y} = \frac{2}{3}h$	$I_{AA} = \frac{bh^3}{36}$ $I_{BB} = \frac{bh^3}{12}$
Círculo	$\bar{x} = \bar{y} = r$	$I_{AA} = \frac{\pi r^4}{4}$ $J_p = \frac{\pi r^4}{2}$
Semicírculo **Semi-circle**	$\bar{y} = \frac{4r}{3\pi}$	$I_{AA} = \frac{r^4}{72\pi}(9\pi^2 - 64)$ $I_{BB} = \frac{\pi r^4}{8}$

Solución. La región consiste en dos rectángulos. Escogiendo los ejes de coordenadas como se indica en la figura 30, los centroides de los rectángulos a y b son

$$\bar{\mathbf{x}}_a = (\bar{x}_a, \bar{y}_a) = (\tfrac{1}{2}, 3)$$

y

$$\bar{\mathbf{x}}_b = (\bar{x}_b, \bar{y}_b) = (\tfrac{7}{2}, \tfrac{1}{2})$$

respectivamente. Denotando las áreas por A_a y A_b respectivamente, tenemos

$$M_x = A_a\bar{y}_a + A_b\bar{y}_b = 1\times 6\times 3 + 5\times 1\times \tfrac{1}{2} = 18 + \tfrac{5}{2} = \tfrac{41}{2} \text{ pulg}^3.$$

Dividiendo por el área total tenemos

$$\bar{y} = \frac{M_x}{A_a + A_b} = \frac{\frac{41}{2}}{6+5} = \frac{41}{22} \text{ pulg.}$$

Por simetría, $\bar{x} = \bar{y} = \frac{41}{22}$ pulgadas.

Para encontrar el momento de inercia respecto a la base —el eje X— notemos primero que los rectángulos tienen, ambos, sus bases sobre el eje X. Por tanto

$$I_x = \frac{1\times 6^3}{3} + \frac{5\times 1^3}{3} = \frac{216+5}{3} = \frac{221}{3} \text{ pulg}^4.$$

18.2 Ejemplo. Encuéntrese el momento de inercia con respecto a la recta que pasa por el centroide y es paralela a la base de la región de la figura 30.

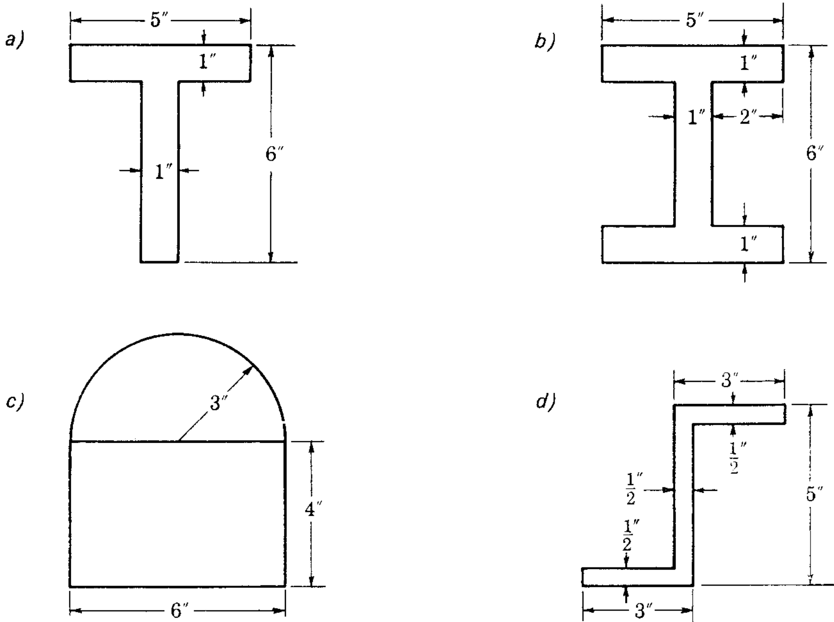
Solución. Podemos usar el teorema de los ejes paralelos (corolario 10.9, pág. 363). Para el rectángulo a la distancia usada en el teorema de los ejes paralelos es $|\bar{y}_a - y|$, y para el rectángulo b la distancia usada es $|\bar{y}_b - y|$. Como el momento de inercia de un rectángulo con respecto a la recta que pasa por el centroide y es paralela a la base es $\frac{1}{12}bh^3$ tenemos para toda la región considerada:

$$I_{AA} = \frac{1\times 6^3}{12} + 1\times 6\times\left(3 - \frac{41}{22}\right)^2 + \frac{5\times 1^3}{12} + 5\times 1\times\left(\frac{1}{2} - \frac{41}{22}\right)^2$$

$$= 18 + \frac{3\,750}{484} + \frac{5}{12} + \frac{4\,500}{484} = \frac{51\,491}{1\,452} \approx 35.46 \text{ pulg}^4.$$

Problemas

1. Encuéntrese el centroide de cada una de las siguientes regiones.



2. Encuéntrense los momentos de inercia con respecto a la base y a una recta paralela que pasa por el centroide de cada una de las regiones del problema 1.

3. Encuéntrese el volumen de la región cilíndrica cuya sección aparece en la figura 31.

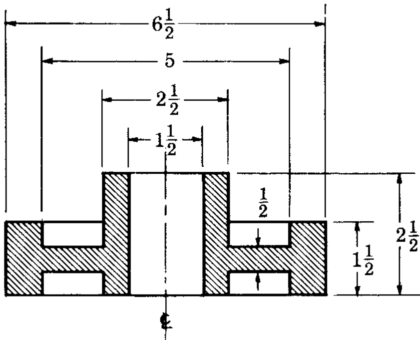


FIGURA 31

19. INTEGRALES MÚLTIPLES

En esta sección, indicaremos cómo puede generalizarse la integral de Riemann a intervalos en R^n. Los casos particulares mas importantes son aquellos en que $n = 1, 2$ y 3. Los teoremas de esta sección son generalizaciones directas de los teoremas de las secciones 2, 3, 4, 5, 6 y 8. En general, no daremos sus pruebas completas, pero indicaremos cuáles son las modificaciones necesarias de las pruebas dadas para sus análogos en el caso bidimensional.

19.1 Definición. *Si* $\mathbf{a} = (a_1, \ldots, a_n)$ *y* $\mathbf{b} = (b_1, \ldots, b_n)$ *con* $a_i \leqslant b_i$ $(i = 1, \ldots, n)$, *entonces el* ***intervalo cerrado*** $[\mathbf{a}, \mathbf{b}]$ en R^n *es el conjunto de todos los puntos* $\mathbf{x} \in \mathrm{R}^n$ *para los cuales*

$$a_i \leqslant x_i \leqslant b_i \quad (i = 1, 2, \ldots, n).$$

El intervalo cerrado $[\mathbf{a}, \mathbf{b}]$ en R^2 es un rectángulo con vértices (a_1, a_2), (b_1, a_2), (b_1, b_2) y (a_1, b_2). Los lados del rectángulo son paralelos a los ejes coordenados. Las longitudes de los lados de $[\mathbf{a}, \mathbf{b}]$ son los números $b_1 - a_1$ y $b_2 - a_2$.

El intervalo cerrado $[\mathbf{a}, \mathbf{b}]$ en R^3 es un paralelepípedo rectangular de vértices (a_1, a_2, a_3), (a_1, b_2, a_3), (a_1, b_2, b_3), (a_1, a_2, b_3), (b_1, a_2, a_3), (b_1, b_2, a_3), (b_1, b_2, b_3) y (b_1, a_2, b_3). De nuevo, los lados del paralelepípedo son paralelos a los ejes de coordenadas y las longitudes de los lados son $b_i - a_i$ $(i = 1, 2, 3)$.

En R, el intervalo $[a, b]$ tiene longitud $b - a$; en R^2, el intervalo $[\mathbf{a}, \mathbf{b}]$ tiene área $(b_1 - a_1)(b_2 - a_2)$; en R^3, el intervalo $[\mathbf{a}, \mathbf{b}]$ tiene volumen $(b_1 - a_1)(b_2 - a_2)(b_3 - a_3)$. Necesitamos un término general que reemplace en R^n a las palabras "longitud" para R, "área" para R^2, y "volumen" para R^3.

19.2 Definición. *El* ***contenido*** *del intervalo* $[\mathbf{a}, \mathbf{b}]$ *en* R^n, *denotado por* $c([\mathbf{a}, \mathbf{b}])$, *es, por definición, el producto*

$$c([\mathbf{a}, \mathbf{b}]) = \prod_{i=1}^{n} (b_i - a_i).$$

En los ejemplos particulares en R, R^2 o R^3, la palabra "contenido" se reemplaza por longitud, área o volumen, respectivamente.

19.3 Definición. *Sea* $[\mathbf{a}, \mathbf{b}] \subset \mathrm{R}^n$ *donde* $\mathbf{a} = (a_1, \ldots, a_n)$ *y* $\mathbf{b} = (b_1, \ldots, b_n)$. *Si para todo* $i = 1, \ldots, n$, *tenemos una partición* $P_i = \{x_i^j \mid j = 0, 1, \ldots, k_i\}$ *de* $[a_i, b_i]$, *donde* $a_i = x_i^0$ *y* $b_i = x_i^{k_i}$, *entonces*

$$P = P_1 \times P_2 \times \ldots \times P_n$$

*se dice que es una **partición** de* [**a**, **b**] *y definimos la **norma** o **malla** de P, lo que escribimos* $|P|$, *por*

$$|P| = \text{máx}\,\{|P_i| \mid i = 1, \ldots, n\}$$

es decir

$$|P| = \text{máx}\,\{x_i^j - x_i^{j-1} \mid j = 1, \ldots, k_i,\ i = 1, \ldots, n\}.$$

Si la partición P_i subdivide a $[a_i, b_i]$ en k_i subintervalos unidimensionales, entonces P subdivide a [**a**, **b**] en $k = k_1 \ldots k_n$ subintervalos n-dimensionales. Los subintervalos de [**a**, **b**] obtenidos por una partición P de [**a**, **b**] pueden enumerarse consecutivamente y denotarse por $\mathscr{R}_i$ con $i = 1, \ldots, k$ donde $k = k_1 \ldots k_n$.

Si f es una función real acotada sobre [**a**, **b**], entonces hay números m y M tales que $m \leqslant f(\mathbf{x}) \leqslant M$ para todo $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$. Como en la sección 2, pág. 314, definimos

$$m_i(f) = \text{ínf}\ \{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\}$$
$$M_i(f) = \sup\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{R}_i\}.$$

Las *sumas inferiores* y las *sumas superiores* se definen por

$$L(f, P) = \sum_{i=1}^{k} m_i(f)\, c(\mathscr{R}_i)$$

y

$$U(f, P) = \sum_{i=1}^{k} M_i(f)\, c(\mathscr{R}_i)$$

respectivamente, donde $c(\mathscr{R}_i)$ es el contenido del subintervalo $\mathscr{R}_i$. Se sigue entonces que

19.4 $$mc([\mathbf{a}, \mathbf{b}]) \leqslant L(f, P) \leqslant U(f, P) \leqslant Mc([\mathbf{a}, \mathbf{b}])$$

para todas las particiones P de [**a**, **b**]. Como en la definición 2.7, pág. 315, definimos la *integral inferior* y la *integral superior* por

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \sup\{L(f, P) \mid P \in \mathscr{P}\}$$

e

$$\overline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \text{ínf}\ \{U(f, P) \mid P \in \mathscr{P}\},$$

respectivamente, donde $\mathscr{P}$ es el conjunto de todas las particiones de [**a**, **b**].

Si la integral inferior y la integral superior de f sobre [**a**, **b**] son iguales, decimos que f es integrable sobre [**a**, **b**].

19.5 Definición. *Una función f sobre* [**a**, **b**] *se dice que es* ***(Riemann) integrable sobre*** [**a**, **b**] *si f está acotada sobre* [**a**, **b**] *e*

$$\underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f.$$

Si f es integrable sobre [**a**, **b**], *entonces la* ***integral definida*** (***de Riemann***) ***de f sobre*** [**a**, **b**], *símbolo* $\int_{\mathbf{a}}^{\mathbf{b}} f$, *está definida por*

$$\int_{\mathbf{a}}^{\mathbf{b}} f = \underline{\int_{\mathbf{a}}^{\mathbf{b}}} f = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} f.$$

La notación $\int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\, d\mathbf{x}$ puede también usarse para representar la integral de f. Si el intervalo [**a**, **b**] es un intervalo en R^2, la integral se llama *integral doble* y si [**a**, **b**] es un intervalo en R^3 la integral se llama *integral triple*. En general, si [**a**, **b**] está en R^n con $n > 1$, la integral se llama *integral múltiple* y si $n = 1$, la integral se llama integral simple. Algunas veces se usan notaciones tales como

$$\iint_{\mathbf{a}}^{\mathbf{b}} f(x_1, x_2)\, dx_1\, dx_2 \quad \text{e} \quad \iiint_{\mathbf{a}}^{\mathbf{b}} f(x_1, x_2, x_3)\, dx_1\, dx_2\, dx_3$$

para denotar a las integrales dobles y triples.

El siguientes resultado es una consecuencia directa de la desigualdad 19.4 y de las definiciones de las integrales inferior y superior.

19.6 Teorema. *Si f es integrable sobre* [**a**, **b**] *y P es una partición cualquiera de* [**a**, **b**], *entonces*

$$L(f, P) \leqslant \int_{\mathbf{a}}^{\mathbf{b}} f \leqslant U(f, P).$$

PRUEBA. Véase el teorema 2.11, pág. 319.

El siguiente teorema muestra que una función acotada es integrable sobre el intervalo [**a**, **b**] en R^n si y sólo si la diferencia entre una suma superior y la correspondiente suma inferior puede hacerse arbitrariamente pequeña.

19.7 Teorema. *Una función acotada f es integrable sobre* [**a**, **b**] *si y sólo si para cada $\varepsilon > 0$ hay una partición P de* [**a**, **b**] *con la propiedad de que* $U(f, P) - L(f, P) < \varepsilon$.

PRUEBA. Véase el teorema 2.14, pág. 320.

Las propiedades básicas de la integral $\int_{\mathbf{a}}^{\mathbf{b}} f$ son las mismas que las de la integral doble (propiedades 3.1-3.8, pág. 322).

Pasamos a continuación al problema de las integrales sobre conjuntos $\mathscr{E}$ más generales sobre R^n.

19.8 Definición. *Si f es una función definida y acotada sobre un conjunto acotado $\mathscr{E}$ en R^n, definimos la función $f_{\mathscr{E}}$ por la regla*

$$f_{\mathscr{E}}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & para\ \mathbf{x} \in \mathscr{E} \\ 0 & para\ \mathbf{x} \in \mathscr{C}\mathscr{E} \end{cases}$$

donde $\mathscr{C}\mathscr{E}$ es el complemento de $\mathscr{E}$ —el conjunto de todos los puntos de R^n no en $\mathscr{E}$.

19.9 Definición. *Si $\mathscr{E}$ es un conjunto acotado en R^n, si $[\mathbf{a}, \mathbf{b}]$ es un intervalo en R^n tal que $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$ y si $\int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}}$ existe, entonces, por definición,*

$$\int_{\mathscr{E}} f = \int_{\mathbf{a}}^{\mathbf{b}} f_{\mathscr{E}}.$$

19.10 Definición. *Si $\mathscr{E}$ es un conjunto acotado en R^n y $[\mathbf{a}, \mathbf{b}]$ es un intervalo en R^n tal que $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$, entonces*

$$\underline{c}(\mathscr{E}) = \underline{\int_{\mathbf{a}}^{\mathbf{b}}} 1_{\mathscr{E}_i} \quad y \quad \bar{c}(\mathscr{E}) = \overline{\int_{\mathbf{a}}^{\mathbf{b}}} 1_{\bar{\mathscr{E}}}$$

se llaman, respectivamente, ***contenido interior*** *y* ***contenido exterior*** *de $\mathscr{E}$. Si $\underline{c}(\mathscr{E}) = \bar{c}(\mathscr{E})$, entonces $1_{\mathscr{E}}$ es integrable sobre $[\mathbf{a}, \mathbf{b}]$ y el valor común se llama* ***contenido*** *de $\mathscr{E}$.*

Si $\mathscr{E}$ tiene contenido se sigue que

$$c(\mathscr{E}) = \int_{\mathbf{a}}^{\mathbf{b}} 1_{\mathscr{E}}.$$

El contenido interior de $\mathscr{E}$, siendo una integral inferior, es el supremo de sumas inferiores. Si P es una partición de $[\mathbf{a}, \mathbf{b}]$ y $\mathscr{R}_j$ denota el j-ésimo subintervalo de P, entonces $m_j(1_{\mathscr{E}_i}) = 0$ para todo subintervalo $\mathscr{R}_j$ que contenga puntos que pertenezcan a $\mathscr{E}_b \cup \mathscr{E}_e$ y por tanto

$$L(1_{\mathscr{E}_i}, P) = \sum_{j=1}^{k} m_j(1_{\mathscr{E}_i})\, c(\mathscr{R}_j)$$

es la suma de los contenidos de aquellos subintervalos de P que son subconjuntos de $\mathscr{E}_i$. Por otra parte, el contenido exterior de $\mathscr{E}$ es una integral superior y, por tanto, es el ínfimo de sumas superiores. Pero $M_j(1_{\bar{\mathscr{E}}}) = 0$ solamente para aquellos subintervalos de P que no contengan ningún punto de $\mathscr{E}$ y por tanto

$$U(1_{\bar{\mathscr{E}}}, P) = \sum_{j=1}^{k} M_j(1_{\bar{\mathscr{E}}})\, c(\mathscr{R}_j)$$

es la suma de los contenidos de aquellos subintervalos de P que contengan puntos de $\bar{\mathscr{E}}$. Así pues, la suma inferior se aproxima al contenido de $\mathscr{E}$ por el contenido de un conjunto inscrito de intervalos y la suma superior se aproxima al contenido de $\mathscr{E}$ por el contenido de un conjunto circunscrito de intervalos.

Podemos demostrar que el contenido tiene propiedades análogas a las propiedades del área. Vemos así que el contenido es no negativo, que el contenido de una parte no es mayor que el contenido del total, y que si un conjunto está dividido en dos partes que no se traslapan, entonces la suma de los contenidos de las partes es igual al contenido del total. Las pruebas de estas propiedades son las mismas que las pruebas de las propiedades correspondientes del área (págs. 333-334).

A continuación demostraremos que el conjunto de las funciones integrables sobre un intervalo cerrado $[\mathbf{a}, \mathbf{b}]$ contiene el conjunto de todas las funciones que son acotadas sobre $[\mathbf{a}, \mathbf{b}]$ para las que el conjunto de puntos de discontinuidad sobre $[\mathbf{a}, \mathbf{b}]$ tiene contenido cero.

19.11 Teorema. *Sea f una función acotada sobre un intervalo $[\mathbf{a}, \mathbf{b}]$ en R^n tal que $\mathscr{E}$, el conjunto de puntos de discontinuidad de f sobre $[\mathbf{a}, \mathbf{b}]$, tenga contenido cero. Entonces $\int_{\mathbf{a}}^{\mathbf{b}} f$ existe.*

PRUEBA. Véase el teorema 5.1, pág. 338. Reemplácese área por contenido y escójase δ de forma tal que $\mathbf{x}^1, \mathbf{x}^2 \in \mathscr{B}$ y $|\mathbf{x}^1 - \mathbf{x}^2| < \sqrt{n}\delta$ implique $|f(\mathbf{x}^1) - f(\mathbf{x}^2)| < \varepsilon$.

Los corolarios 5.2 a 5.6, págs. 339-340, se generalizan imnediatamente a R^n al igual que todos los resultados de la sección 6.

Es posible introducir regiones en R^n y discutir integrales sobre regiones. No haremos esto; en lugar de ello estableceremos la relación fundamental entre las integrales múltiples y las integrales iteradas.

19.12 Teorema. *Si $\int_{\mathbf{a}}^{\mathbf{b}} f$ existe, y si para cada $y \in [\mathbf{c}, \mathbf{d}]$, donde $[\mathbf{c}, \mathbf{d}]$ es la proyección de $[\mathbf{a}, \mathbf{b}]$ paralela al eje X_n, $F(\mathbf{y}) = \int_{a_n}^{b_n} f(\mathbf{y}, x_n)\, dx_n$ existe,*

entonces $\int_{\mathbf{c}}^{\mathbf{d}} \left\{ \int_{a_n}^{b_n} f(\mathbf{y}, x_n)\, dx_n \right\} d\mathbf{y} = \int_{\mathbf{c}}^{\mathbf{d}} F(\mathbf{y})\, d\mathbf{y}$ *existe e*

$$\int_{\mathbf{a}}^{\mathbf{b}} f = \int_{\mathbf{c}}^{\mathbf{d}} \left\{ \int_{a_n}^{b_n} f(\mathbf{y}, x_n)\, dx_n \right\} d\mathbf{y}.$$

PRUEBA. (Véase el teorema 8.1, pág. 000.) Probaremos primero que para cualquier partición $P = P_1 \times P_2 \times \ldots \times P_n$ de $[\mathbf{a}, \mathbf{b}]$

$$L(f, P) \leqslant L(F, P') \leqslant U(F, P') \leqslant U(f, P)$$

donde $P' = P_1 \times P_2 \times \ldots \times P_{n-1}$ es la partición de $[\mathbf{c}, \mathbf{d}]$ inducida por la partición P de $[\mathbf{a}, \mathbf{b}]$. La existencia de la integral iterada y la igualdad de la integral múltiple y la integral iterada se sigue de esta desigualdad.

Sea $P = P_1 \times P_2 \times \ldots \times P_n$ una partición cualquiera de $[\mathbf{a}, \mathbf{b}]$. P induce una partición $P' = P_1 \times P_2 \times \ldots \times P_{n-1}$ de $[\mathbf{c}, \mathbf{d}]$ en $k = k_1 k_2 \ldots k_{n-1}$ subintervalos $\mathscr{R}_i$. Sean

$$m_{ij}(f) = \inf \{ f(\mathbf{y}, x_n) \mid \mathbf{y} \in \mathscr{R}_i,\ x_n \in [x_n^{j-1}, x_n^j] \}$$

$$m_j(f; \mathbf{y}) = \inf \{ f(\mathbf{y}, x_n) \mid x_n \in [x_n^{j-1}, x_n^j] \}$$

$$M_{ij}(f) = \sup \{ f(\mathbf{y}, x_n) \mid \mathbf{y} \in \mathscr{R}_i,\ x_n \in [x_n^{j-1}, x_n^j] \}$$

$$M_j(f; \mathbf{y}) = \sup \{ f(\mathbf{y}, x_n) \mid x_n \in [x_n^{j-1}, x_n^j] \}.$$

Entonces, para cualquier $\mathbf{y} \in \mathscr{R}_i$,

$$\begin{aligned} \sum_{j=1}^{k_n} m_{ij}(f)(x_n^j - x_n^{j-1}) &\leqslant \sum_{j=1}^{k_n} m_j(f; \mathbf{y})(x_n^j - x_n^{j-1}) = L(f, P_n; \mathbf{y}) \\ &\leqslant \int_{a_n}^{b_n} f(\mathbf{y}, x_n)\, dx_n \\ &\leqslant U(f, P_n; \mathbf{y}) = \sum_{j=1}^{k_n} M_j(f; \mathbf{y})(x_n^j - x_n^{j-1}) \\ &\leqslant \sum_{j=1}^{k_n} M_{ij}(f)(x_n^j - x_n^{j-1}). \end{aligned}$$

Es decir, para toda $\mathbf{y} \in \mathscr{R}_i$

$$\sum_{j=1}^{k_n} m_{ij}(f)(x_n^j - x_n^{j-1}) \leqslant F(\mathbf{y}) \leqslant \sum_{j=1}^{k_n} M_{ij}(f)(x_n^j - x_n^{j-1}).$$

Como la anterior desigualdad se verifica para toda $\mathbf{y} \in \mathscr{R}_i$, de ello se sigue que

19.13 $$\sum_{j=1}^{k_n} m_{ij}(f)(x_n^j - x_n^{j-1}) \leqslant m_i(F) \leqslant M_i(F) \leqslant \sum_{j=1}^{k_n} M_{ij}(f)(x_n^j - x_n^{j-1})$$

donde

$$m_i(F) = \inf \{F(\mathbf{y}) \mid \mathbf{y} \in \mathscr{R}_i\}$$

y

$$M_i(F) = \sup \{F(\mathbf{y}) \mid \mathbf{y} \in \mathscr{R}_i\}.$$

Multiplicando la desigualdad 19.13 por $c(\mathscr{R}_i)$ y sumando desde $i = 1$ hasta $i = k$, obtenemos

$$\begin{aligned} L(f, P) &= \sum_{i=1}^{k} \sum_{j=1}^{k_n} m_{ij}(f)\,(x_n^j - x_n^{j-1})\, c(\mathscr{R}_i) \\ &\leqslant \sum_{i=1}^{k} m_i(F)\, c(\mathscr{R}_i) = L(F, P') \\ &\leqslant U(F, P') = \sum_{i=1}^{k} M_i(F)\, c(\mathscr{R}_i) \\ &\leqslant \sum_{i=1}^{k} \sum_{j=1}^{k_n} M_{ij}(f)\,(x_n^j - x_n^{j-1})\, c(\mathscr{R}_i) = U(f, P), \end{aligned}$$

o bien

19.14 $$L(f, P) \leqslant L(F, P') \leqslant U(F, P') \leqslant U(f, P).$$

Como $\int_{\mathbf{a}}^{\mathbf{b}} f$ existe, dada una $\varepsilon > 0$ existe una partición P de $[\mathbf{a}, \mathbf{b}]$ tal que $U(f, P) - L(f, P) < \varepsilon$. De 19.14 se concluye que

$$U(F, P') - L(F, P') \leqslant U(f, P) - L(f, P) < \varepsilon$$

y por tanto $\int_{\mathbf{c}}^{\mathbf{d}} F(\mathbf{y})\, d\mathbf{y}$ existe. Por otra parte $L(f, P) \leqslant \int_{\mathbf{a}}^{\mathbf{b}} f \leqslant U(f, P)$ y

$$L(F, P') \leqslant \int_{\mathbf{c}}^{\mathbf{d}} F(\mathbf{y})\, d\mathbf{y} \leqslant U(F, P')$$

de modo que

$$\left| \int_{\mathbf{a}}^{\mathbf{b}} f - \int_{\mathbf{c}}^{\mathbf{d}} F \right| \leqslant U(f, P) - L(f, P) < \varepsilon.$$

Por tanto

$$\int_{\mathbf{a}}^{\mathbf{b}} f = \int_{\mathbf{c}}^{\mathbf{d}} F = \int_{\mathbf{c}}^{\mathbf{d}} \left\{ \int_{a_n}^{b_n} f(\mathbf{y}, x_n)\, dx_n \right\} d\mathbf{y}.$$

Y esto completa la prueba.

Damos a continuación condiciones suficientes para asegurar la existencia de las integrales del teorema 19.12.

19.15 Corolario. *Sea f una función acotada sobre un intervalo* $[\mathbf{a}, \mathbf{b}]$ *en* R^n *y continua sobre* $[\mathbf{a}, \mathbf{b}]$, *excepto sobre un conjunto* $\mathscr{E}$ *de contenido cero, con la propiedad de que toda recta paralela al eje* X_n *intersecta con, cuando más, en un número finito de puntos. Entonces*

$$\int_{\mathbf{a}}^{\mathbf{b}} f = \int_{\mathbf{c}}^{\mathbf{d}} \left\{ \int_{a_n}^{b_n} f(\mathbf{y}, x_n)\, dx_n \right\} d\mathbf{y}$$

donde $[\mathbf{c}, \mathbf{d}]$ *es la proyección de* $[\mathbf{a}, \mathbf{b}]$ *paralela al eje* X_n.

PRUEBA. Por el teorema 19.11, $\int_{\mathbf{a}}^{\mathbf{b}} f$ existe. Por otra parte, toda recta paralela al eje X_n que pasa por $[\mathbf{c}, \mathbf{d}]$ intersecta a $\mathscr{E}$ en, cuando más, un número finito de puntos de modo que f tiene cuando más un número finito de discontinuidades sobre tal recta. Luego

$$F(\mathbf{y}) = \int_{a_n}^{b_n} f(\mathbf{y}, x_n)\, dx_n$$

existe para toda $\mathbf{y} \in [\mathbf{c}, \mathbf{d}]$. Por tanto, por el teorema 19.12,

$$\int_{\mathbf{a}}^{\mathbf{b}} f = \int_{\mathbf{c}}^{\mathbf{d}} F = \int_{\mathbf{c}}^{\mathbf{d}} \left\{ \int_{a_n}^{b_n} f(\mathbf{y}, x_n)\, dx_n \right\} d\mathbf{y}.$$

19.16 Corolario. *Sea f una función acotada sobre un intervalo* $[\mathbf{a}, \mathbf{b}]$ *en* R^n *y continua sobre* $[\mathbf{a}, \mathbf{b}]$, *excepto sobre un conjunto* $\mathscr{E}$ *de contenido cero, con la propiedad de que toda recta paralela al eje* X_i $(i = 1, \ldots, n)$ *intersecta a* $\mathscr{E}$ *en, cuando más, un número finito de puntos. Entonces*

$$\int_{\mathbf{a}}^{\mathbf{b}} f = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \cdots, x_n)\, dx_n \cdots dx_1 .$$

PRUEBA. El corolario sigue del teorema 19.12 por inducción. Sea $\mathscr{S}$ el conjunto de todos los enteros positivos n para los que el corolario se verifica. Claramente $1 \in \mathscr{S}$. Supongamos que $m \in \mathscr{S}$, es decir, que

$$\int_{\mathbf{c}}^{\mathbf{d}} F(\mathbf{y})\, d\mathbf{y} = \int_{a_1}^{b_1} \cdots \int_{a_m}^{b_m} F(x_1, \cdots, x_m)\, dx_m \cdots dx_1 .$$

Consideremos ahora $n = m+1$. Si toda recta paralela al eje X_{m+1} intersecta con $\mathscr{E}$ en cuando más un número finito de puntos, entonces f tiene cuando más un número finito de discontinuidades sobre una tal recta. Luego

$$F(\mathbf{y}) = F(x_1, \cdots, x_m) = \int_{a_{m+1}}^{b_{m+1}} f(x_1, \cdots, x_m, x_{m+1})\, dx_{m+1}$$

existe para todo $\mathbf{y} \in [\mathbf{c}, \mathbf{d}]$, y, por el teorema 19.12,

$$\int_{(a_1, \ldots, a_{m+1})}^{(b_1, \ldots, b_{m+1})} f = \int_{(a_1, \ldots, a_m)}^{(b_1, \ldots, b_m)} \left\{ \int_{a_{m+1}}^{b_{m+1}} f(x_1, \cdots, x_m, x_{m+1})\, dx_{m+1} \right\} d(x_1, \cdots, x_m)$$

$$= \int_{(a_1, \ldots, a_m)}^{(b_1, \ldots, b_m)} F(x_1, \cdots, x_m)\, d(x_1, \cdots, x_m)$$

$$= \int_{a_1}^{b_1} \cdots \int_{a_m}^{b_m} F(x_1, \cdots, x_m)\, dx_m \cdots dx_1$$

$$= \int_{a_1}^{b_1} \cdots \int_{a_m}^{b_m} \int_{a_{m+1}}^{b_{m+1}} f(x_1, \cdots, x_m, x_{m+1})\, dx_{m+1}\, dx_m \cdots dx_1$$

y $m+1 \in \mathscr{S}$. De donde el corolario se verifica para todos los enteros positivos n.

20. RESUMEN

En este capítulo extendimos la teoría de la integración desde las integrales simples hasta las integrales múltiples. Vimos que después de definir los intervalos en R^2, R^3 o, en general, R^n, la teoría de la integral múltiple es exactamente paralela a la de la integral simple. En realidad, si se desarrolla la teoría para intervalos en R^n, obtenemos la integral simple como el caso particular en que $n = 1$. Deben introducirse algunas nuevas consideraciones cuando deseamos extender el concepto de la integral a conjuntos acotados más generales de R^n. La evaluación de integrales múltiples puede reducirse en muchos casos a la evaluación de integrales iteradas, es decir, a la evaluación de integrales simples sucesivas. Las integrales simples sucesivas pueden evaluarse a veces usando el segundo teorema fundamental del cálculo: si F es una función tal que $F' = f$ sobre $[a, b]$ entonces $\int_b^a f = F(b) - F(a)$. Si el teorema fundamental del cálculo se prueba que es inaplicable, entonces puede usarse la integración numérica.

Problemas

1. Encuéntrese el área de las regiones limitadas por los siguientes conjuntos de curvas.

a) $y^2 = 2+2x$, $y^2 = 2-4x$
b) $x = 0$, $y = 0$, $x = 4$, $y = e^x$
c) $\sqrt{x}+\sqrt{y} = \sqrt{a}$, $x = 0$, $y = 0$
d) el rizo de la hoja de Descartes $y^2(a+x) = x^2(3a-x)$.

2. Encuéntrese el centroide de las regiones limitadas por los siguientes conjuntos de curvas.

a) $x = 0$, $x = 1$, $y = \operatorname{senh} x$, $y = \cosh x$
b) $y = 2-x^2$, $y = x$
c) $x = y^2$, $y = x-2$
d) $y = \operatorname{sen} \pi x$, $y = x^2-x$.

3. Encuéntrese el momento de inercia de las regiones limitadas por las siguientes curvas con respecto a la recta dada.

a) $y = 2-x^2$, $y = x$ con respecto a $x = \sqrt{2}$
b) $y = 2-x^2$, $y = x$ con respecto a $y = 1$.

4. Encuéntrese el momento de inercia con respecto al eje Y de la región limitada superiormente por la serpentina $(a^2+x^2)y = 2a^2x$, en la parte inferior por el eje X, a la izquierda por la recta vertical que pasa por el punto máximo, y a la derecha por la recta vertical que pasa por el punto de inflexión con abscisa positiva.

5. La fuerza total ejercida por un fluido sobre una región plana está definida como la integral sobre la región de la presión del fluido donde la presión en un punto es el producto del peso por unidad de volumen del fluido por la profundidad del punto respecto a la superficie del fluido. Demuéstrese que para una región en un plano vertical la fuerza es igual al producto del área de la región por la presión en el centroide de la región.

6. Encuéntrese la fuerza total debida a la presión del agua sobre cada una de las siguientes superficies verticales. Todas las distancias están medidas en pies. El peso del agua es aproximadamente de 62.5 libras por pie cúbico.

a) limitada por la circunferencia $x^2+y^2 = 25$; el nivel del agua sobre el eje X
b) limitada por la parábola $y = x^2-4$ y el eje X; el nivel del agua sobre el eje X
c) limitada por la elipse $16x^2+25y^2 = 400$; el nivel del agua sobre el eje X
d) limitada por la elipse $16x^2+25y^2 = 400$; el nivel del agua sobre la recta $y = 4$.





Funciones de conjunto e integrales múltiples

1. INTRODUCCIÓN

En la mayoría de los casos, las funciones que hasta ahora hemos considerado tenían conjuntos de puntos de un espacio euclidiano R^n como dominio. Ahora consideraremos funciones con una familia (conjunto) de conjuntos como dominio; tales funciones se llaman ***funciones de conjunto.*** En particular, nos ocuparemos de funciones de conjunto con una familia de conjuntos en R^n como dominio y con rango en R. Ya nos hemos encontrado con algunas funciones de tal tipo. Por ejemplo, si $\mathfrak{R}$ es la familia de subconjuntos de R^2 que tienen área, entonces la función A con regla de correspondencia

$$A(\mathscr{E}) = \int_{\mathscr{E}} 1 \quad \text{para} \quad \mathscr{E} \in \mathfrak{R}$$

es una función real de conjunto (una función de conjunto valuada en el campo real).

A las funciones cuyo dominio es un conjunto de puntos en R^n las llamaremos ***funciones de punto*** para distinguirlas de las funciones de conjunto.

2. ANILLOS DE CONJUNTOS

Recuérdese que si $\mathscr{A} \subset \mathscr{S}$, el complemento de $\mathscr{A}$ con respecto a $\mathscr{S}$, denotado por $\mathscr{C}_{\mathscr{S}}\mathscr{A}$, es el conjunto de todos los elementos $x \in \mathscr{S}$ tales que $x \notin \mathscr{A}$. Cuando en una discusión determinada $\mathscr{S}$ es fijo, denotamos al complemento de $\mathscr{A}$ con respecto a $\mathscr{S}$ por $\mathscr{C}\mathscr{A}$ y hablamos simplemente del complemento de $\mathscr{A}$. Si $\mathscr{A}$ y $\mathscr{B}$ son dos subconjuntos de un conjunto $\mathscr{S}$, el conjunto diferencia $\mathscr{A} - \mathscr{B}$ es el conjunto de todos los elementos $x \in \mathscr{A}$ tales que $x \notin \mathscr{B}$, es decir, $\mathscr{A} - \mathscr{B} = \mathscr{A} \cap \mathscr{C}\mathscr{B}$. *No* exigimos que $\mathscr{B}$ sea un subconjunto de $\mathscr{A}$ para que la diferencia $\mathscr{A} - \mathscr{B}$ esté definida.

2.1 Definición. *Sea $\mathscr{S}$ un conjunto, y $\mathfrak{R}$ una familia no vacía de subconjuntos de $\mathscr{S}$. La familia $\mathfrak{R}$ se llama* ***anillo*** *(de conjuntos), o* ***anillo booleano****, si $\mathscr{E}$, $\mathscr{F} \in \mathfrak{R}$ implica*

$$\mathscr{E} \cup \mathscr{F} \in \mathfrak{R} \quad y \quad \mathscr{E} - \mathscr{F} \in \mathfrak{R}.$$

Como consecuencia inmediata de la definición de un anillo, si $\mathscr{E}$, $\mathscr{F} \in \mathfrak{R}$, entonces $\mathscr{E} \cap \mathscr{F} \in \mathfrak{R}$ ya que $\mathscr{E} \cap \mathscr{F} = \mathscr{E} - (\mathscr{E} - \mathscr{F})$. Además, todo anillo contiene el conjunto vacío, $\varnothing$, ya que $\mathscr{E} - \mathscr{E} = \varnothing$ para todo $\mathscr{E} \in \mathfrak{R}$.

2.2 Ejemplo. Demuéstrese que si $\mathscr{G}$ es un conjunto acotado en R^2 que tiene área y $\mathfrak{R}$, es la familia de subconjuntos de $\mathscr{G}$ que tienen área, entonces $\mathfrak{R}$ es un anillo.

SOLUCIÓN. $\mathfrak{R}$ no es vacío ya que $\mathscr{G} \in \mathfrak{R}$. Sean $\mathscr{E}$ y $\mathscr{F}$ elementos de $\mathfrak{R}$. Entonces $\mathscr{E} \cup \mathscr{F} \subset \mathscr{G}$ y por el teorema 4.12, pág. 334, $\mathscr{E} \cup \mathscr{F}$ tiene área y por tanto pertenece a $\mathfrak{R}$. Además, $\mathscr{E} - \mathscr{F} \in \mathfrak{R}$ ya que $\mathscr{E} - \mathscr{F} \subset \mathscr{G}$ y, por el teorema 4.15, pág. 336, $\mathscr{E} - \mathscr{F}$ tiene área.

El ejemplo 2.2 se generaliza inmediatamente a la familia de subconjuntos de un conjunto acotado que tienen contenido en R^n. En este capítulo nos ocuparemos exclusivamente de funciones de conjunto reales definidas sobre un anillo de conjuntos con contenido en R^n.

En álgebra moderna un *anillo* es un conjunto R con dos operaciones, adición y multiplicación, que satisfacen las propiedades A_1 a A_5, M_1, M_3, y D enumeradas en la página 37. Si la ley conmutativa para la multiplicación, M_2, también se verifica, el anillo se llama *anillo conmutativo.*

Si consideramos la diferencia simétrica de conjuntos definida por

$\mathscr{E} \Delta \mathscr{F} = (\mathscr{E} - \mathscr{F}) \cup (\mathscr{F} - \mathscr{E})$ como operación de adición y la intersección de conjuntos como la operación de multiplicación, entonces es fácil probar que un anillo de conjuntos es un anillo, y precisamente un anillo conmutativo, en el sentido del álgebra. Como para todo $\mathscr{E} \in \mathfrak{R}$, $\varnothing \Delta \mathscr{E} = \mathscr{E}$, el conjunto vacío juega el papel de 0 en el anillo de conjuntos. Si $\mathscr{X} = \bigcup_{\mathscr{E} \in \mathfrak{R}} \mathscr{E}$ está en $\mathfrak{R}$, entonces para todo $\mathscr{E} \in \mathfrak{R}$, $\mathscr{X} \cap \mathscr{E} = \mathscr{E}$ y $\mathscr{X}$ desempeña el papel de 1 en el anillo de conjuntos. Un anillo de conjuntos se llama *álgebra* o *álgebra booleana* si $\mathscr{X} \in \mathfrak{R}$.

3. FUNCIONES DE CONJUNTO

3.1 Definición. *Una función de conjunto F sobre un anillo $\mathfrak{S}$ de conjuntos en* R^n *se dice que es* ***finitamente aditiva*** *si*

$$F(\mathscr{E} \cup \mathscr{F}) = F(\mathscr{E}) + F(\mathscr{F})$$

siempre que $\mathscr{E}$, $\mathscr{F} \in \mathfrak{S}$ y $c(\mathscr{E} \cap \mathscr{F}) = 0$.

Por ejemplo, la función contenido es finitamente aditiva sobre un anillo de conjuntos que tengan contenido.

Nota. En lo que falta de este capítulo usaremos el término "aditiva" para indicar "finitamente aditiva".

3.2 Ejemplo. Pruébese que si f es integrable sobre un intervalo $[\mathbf{a}, \mathbf{b}] \subset \mathrm{R}^2$, entonces la función F definida sobre el anillo $\mathfrak{R}$ de subconjuntos de $[\mathbf{a}, \mathbf{b}]$ que tienen área por la regla

$$F(\mathscr{E}) = \int_{\mathscr{E}} f \qquad \mathscr{E} \in \mathfrak{R}$$

es finitamente aditiva.

Solución. Primero demostraremos que si f es integrable sobre $[\mathbf{a}, \mathbf{b}]$, entonces F está definida sobre el anillo $\mathfrak{R}$. Si $\mathscr{E} \in \mathfrak{R}$, entonces $\mathscr{E}$ tiene área. Como $[\mathbf{a}, \mathbf{b}]$ tiene también área y f es integrable sobre $[\mathbf{a}, \mathbf{b}]$, por el corolario 6.14, pág. 345, podemos concluir que f es integrable sobre $\mathscr{E}$. Si $\mathscr{E}$, $\mathscr{F}$ son subconjuntos de $[\mathbf{a}, \mathbf{b}]$ que tienen área y $A(\mathscr{E} \cap \mathscr{F}) = 0$, entonces, según el corolario 6.13, pág. 345,

$$F(\mathscr{E} \cup \mathscr{F}) = \int_{\mathscr{E} \cup \mathscr{F}} f = \int_{\mathscr{E}} f + \int_{\mathscr{F}} f = F(\mathscr{E}) + F(\mathscr{F}).$$

El ejemplo 3.2 se generaliza inmediatamente a conjuntos que tienen contenido en R^n.

Definimos la suma de dos funciones de conjunto reales F y G como la

función de conjunto $F+G$ de dominio $\mathfrak{D}_F \cap \mathfrak{D}_G$ y regla de correspondencia

$$[F+G](\mathscr{E}) = F(\mathscr{E})+G(\mathscr{E}) \qquad \mathscr{E}\in\mathfrak{D}_F \cap \mathfrak{D}_G .$$

Análogamente, definimos la diferencia $F-G$.

Si F y G son funciones de conjunto finitamente aditivas sobre un anillo $\mathfrak{G}$, entonces $F+G$ y $F-G$ son finitamente aditivas sobre $\mathfrak{G}$. Por ejemplo, si $\mathscr{E}$, $\mathscr{F}\in\mathfrak{G}$ y $c(\mathscr{E}\cap\mathscr{F}) = 0$, entonces

$$\begin{aligned}[F-G](\mathscr{E}\cup\mathscr{F}) &= F(\mathscr{E}\cup\mathscr{F})-G(\mathscr{E}\cup\mathscr{F})\\ &= F(\mathscr{E})+F(\mathscr{F})-G(\mathscr{E})-G(\mathscr{F})\\ &= [F-G](\mathscr{E})+[F-G](\mathscr{F}).\end{aligned}$$

3.3 Definición. *Una función de conjunto aditiva F se dice que es* ***monótona*** *(no decreciente) si sus valores son todos no negativos.*

Se sigue fácilmente de la definición 3.3 que si F es una función de conjunto aditiva monótona definida sobre un anillo $\mathfrak{R}$ y $\mathscr{E}$, $\mathscr{F}\in\mathfrak{R}$ con $\mathscr{E}\subset\mathscr{F}$, entonces $F(\mathscr{E}) \leqslant F(\mathscr{F})$; en efecto, $\mathscr{E}$, $\mathscr{F}\in\mathfrak{R}$ y $\mathscr{E}\subset\mathscr{F}$ implica $\mathscr{F}-\mathscr{E}\in\mathfrak{R}$ y

$$F(\mathscr{F}) = F(\mathscr{E}\cup(\mathscr{F}-\mathscr{E})) = F(\mathscr{E})+F(\mathscr{F}-\mathscr{E}) \geqslant F(\mathscr{E}).$$

La función contenido es un ejemplo de una función monótona. Además, si la función f del ejemplo 3.2 tiene sólo valores positivos sobre $[\mathbf{a}, \mathbf{b}]$, entonces, para $\mathscr{E}\in\mathfrak{R}$, por la propiedad 6.7, pág. 342,

$$F(\mathscr{E}) = \int_{\mathscr{E}} f \geqslant \int_{\mathscr{E}} 0 = 0$$

y F es monótona.

Definimos a continuación los conceptos de límite y derivada. Siempre que hablamos del límite de una función de conjunto F en un punto $\mathbf{x}_0$ suponemos que $\mathbf{x}_0$ es un punto de acumulación del dominio de F. Un punto $\mathbf{x}_0$ se dice que es un ***punto de acumulación*** de una familia $\mathfrak{R}$ de subconjuntos de R^n si para todo número $\delta > 0$ existe un conjunto $\mathscr{E}\in\mathfrak{R}$ de contenido positivo que contiene a $\mathbf{x}_0$ y tiene diámetro $d(\mathscr{E}) < \delta$, donde $d(\mathscr{E}) = \sup\{|\mathbf{x}-\mathbf{y}| \mid \mathbf{x}, \mathbf{y}\in\mathscr{E}\}$.

3.4 Definición. *Sea F una función real de conjunto definida sobre la familia $\mathfrak{D}_F$ de conjuntos de* R^n. *La función F se dice que tiene el* ***límite*** *b en* $\mathbf{x}_0$, *lo que se escribe*: $\lim_{\mathbf{x}_0} F = b$ *o* $\lim_{\mathscr{E}\to\mathbf{x}_0} F(\mathscr{E}) = b$, *si* $\mathbf{x}_0$ *es un punto de acumulación de la familia* $\mathfrak{D}_F$ *y si para cada* $\varepsilon > 0$ *existe una* $\delta > 0$ *tal que*

$$|F(\mathscr{E})-b| < \varepsilon$$

siempre que $\mathscr{E}\in\mathfrak{D}_F$ *es un conjunto de contenido positivo tal que* $\mathbf{x}_0\in\mathscr{E}$ *y* $d(\mathscr{E}) < \delta$.

Si $\lim_{\mathbf{x}} F$ existe para todo $\mathbf{x}$ en un conjunto $\mathscr{S}$, entonces obtenemos una función de punto f definida sobre $\mathscr{S}$ con regla de correspondencia $f(\mathbf{x}) = \lim_{\mathbf{x}} F$.

Puede probarse fácilmente que si $\lim_{\mathbf{x}_0} F$ y $\lim_{\mathbf{x}_0} G$ existen y $\mathbf{x}_0$ es un punto de acumulación de $\mathfrak{D}_F \cap \mathfrak{D}_G$, entonces

$$\lim_{\mathbf{x}_0} [F+G] = \lim_{\mathbf{x}_0} F + \lim_{\mathbf{x}_0} G$$

y

$$\lim_{\mathbf{x}_0} [F-G] = \lim_{\mathbf{x}_0} F - \lim_{\mathbf{x}_0} G .$$

Si definimos el producto y el cociente de dos funciones de conjunto reales como las funciones FG y F/G con reglas de correspondencia

$$[FG](\mathscr{E}) = F(\mathscr{E})\, G(\mathscr{E}) \quad \text{y} \quad \left[\frac{F}{G}\right](\mathscr{E}) = \frac{F(\mathscr{E})}{G(\mathscr{E})}$$

y dominios $\mathfrak{D}_{FG} = \mathfrak{D}_F \cap \mathfrak{D}_G$ y $\mathfrak{D}_{F/G} = \{\mathscr{E} \in \mathfrak{D}_F \cap \mathfrak{D}_G \mid G(\mathscr{E}) \neq 0\}$, respectivamente, entonces los teoremas sobre producto y cociente de límites son válidos también para estas funciones.

3.5 Definición. *Sea $\mathscr{S}$ un conjunto de puntos de acumulación del dominio $\mathfrak{D}_F$ de una función de conjunto F. Una función de punto f se dice que es el* ***límite uniforme*** *de F sobre $\mathscr{S}$ si para todo $\varepsilon > 0$ existe un número $\delta > 0$ tal que*

$$|F(\mathscr{E}) - f(\mathbf{x})| < \varepsilon$$

siempre que $\mathbf{x} \in \mathscr{S}$ y $\mathscr{E} \in \mathfrak{D}_F$ es un conjunto con contenido positivo tal que $\mathbf{x} \in \mathscr{E}$ y $d(\mathscr{E}) < \delta$.

3.6 Definición. *Una función de conjunto F definida sobre una familia de conjuntos en R^n se dice que es* ***diferenciable*** *en el punto $\mathbf{x}$ y que tiene* ***derivada*** *$[DF](\mathbf{x})$ si el límite*

$$[DF](\mathbf{x}) = \lim_{\mathscr{E} \to \mathbf{x}} \frac{F(\mathscr{E})}{c(\mathscr{E})}$$

existe. La función de conjunto F es diferenciable sobre un conjunto $\mathscr{S}$ si F es diferenciable en cada punto de $\mathscr{S}$ y F es uniformemente diferenciable sobre $\mathscr{S}$ si DF es el límite uniforme de F/c sobre $\mathscr{S}$. La función de punto DF se llama la derivada de F.

Se prueba fácilmente que las reglas para la derivada de una suma y una diferencia de dos funciones de conjunto reales toman la misma forma que las reglas dadas para el caso de funciones reales de una variable real.

4. EL TEOREMA FUNDAMENTAL DEL CÁLCULO

En esta sección probaremos dos teoremas que son análogos al primero y segundo teoremas fundamentales del cálculo. Para funciones reales de variable real el primer teorema fundamental del cálculo rezaba: si f es continua sobre un intervalo $\mathscr{J}$, entonces

$$D_x\left[\int_a^x f(t)\,dt\right] = f(x)$$

para $a, x \in \mathscr{J}$ cualesquiera. El segundo teorema fundamental nos dice: si F tiene una derivada continua sobre un intervalo $\mathscr{J}$, entonces para todo $a, b \in \mathscr{J}$

$$\int_a^b D_x[F(x)]\,dx = F(b) - F(a).$$

Antes de considerar los análogos de los teoremas fundamentales del cálculo introducimos las nociones de distancia entre un punto y un conjunto y de distancia entre dos conjuntos. La distancia de un punto $\mathbf{x}$ a un conjunto $\mathscr{A}$ en R^n, denotada por $d(\mathbf{x}, \mathscr{A})$, se define como sigue:

$$d(\mathbf{x}, \mathscr{A}) = \text{ínf}\,\{|\mathbf{x}-\mathbf{y}| \mid \mathbf{y} \in \mathscr{A}\}.$$

Claramente, $d(\mathbf{x}, \mathscr{A}) \geqslant 0$ y si $\mathbf{x} \in \mathscr{A}$, entonces $d(\mathbf{x}, \mathscr{A}) = 0$. Probaremos ahora que si $\mathscr{A}$ es cerrado y $\mathbf{x} \notin \mathscr{A}$, entonces $d(\mathbf{x}, \mathscr{A}) > 0$. Como $\mathbf{x} \in \mathscr{C}\mathscr{A}$ y $\mathscr{C}\mathscr{A}$ es abierto, existe una vecindad $\mathscr{S}(\mathbf{x}; \varepsilon)$ de $\mathbf{x}$ que está contenida en $\mathscr{C}\mathscr{A}$. Luego $d(\mathbf{x}, \mathscr{A}) \geqslant \varepsilon > 0$. Si $\mathscr{A}$ no es cerrado, podemos tener $d(\mathbf{x}, \mathscr{A}) = 0$, aun cuando $x \notin \mathscr{A}$. Por ejemplo, si $\mathscr{A}$ es el disco abierto unitario $\{\mathbf{x} \mid |\mathbf{x}| < 1\}$ en R^2 y $\mathbf{x} = (1, 0)$, entonces $d(\mathbf{x}, \mathscr{A}) = 0$. En general, si $\mathbf{x}$ es un punto frontera de $\mathscr{A}$, entonces $d(\mathbf{x}, \mathscr{A}) = 0$. Para un conjunto $\mathscr{A}$ fijo en R^n, la función $d_{\mathscr{A}}$ definida por $d_{\mathscr{A}}(\mathbf{x}) = d(\mathbf{x}, \mathscr{A})$ es continua sobre R^n. La continuidad de $d_{\mathscr{A}}$ se sigue de la desigualdad del triángulo:

$$|\mathbf{x}-\mathbf{z}| \leqslant |\mathbf{x}-\mathbf{y}| + |\mathbf{y}-\mathbf{z}|.$$

Así pues, para cualesquier $\mathbf{x}, \mathbf{y} \in R^n$,

$$\begin{aligned} d_{\mathscr{A}}(\mathbf{x}) &= \text{ínf}\,\{|\mathbf{x}-\mathbf{z}| \mid \mathbf{z} \in \mathscr{A}\} \leqslant \text{ínf}\,\{|\mathbf{x}-\mathbf{y}| + |\mathbf{y}-\mathbf{z}| \mid \mathbf{z} \in \mathscr{A}\} \\ &= |\mathbf{x}-\mathbf{y}| + \text{ínf}\,\{|\mathbf{y}-\mathbf{z}| \mid \mathbf{z} \in \mathscr{A}\} = |\mathbf{x}-\mathbf{y}| + d_{\mathscr{A}}(\mathbf{y}). \end{aligned}$$

Análogamente, $d_{\mathscr{A}}(\mathbf{y}) \leqslant |\mathbf{x}-\mathbf{y}| + d_{\mathscr{A}}(\mathbf{x})$. Por tanto,

$$|d_{\mathscr{A}}(\mathbf{x}) - d_{\mathscr{A}}(\mathbf{y})| \leqslant |\mathbf{x}-\mathbf{y}|$$

de modo que $|d_{\mathscr{A}}(\mathbf{x}) - d_{\mathscr{A}}(\mathbf{y})| < \varepsilon$ siempre que $|\mathbf{x}-\mathbf{y}| < \delta = \varepsilon$.

Si $\mathscr{A}$ y $\mathscr{B}$ son conjuntos en R^n, la distancia entre $\mathscr{A}$ y $\mathscr{B}$, denotada por $d(\mathscr{A}, \mathscr{B})$, se define como sigue:

$$d(\mathscr{A}, \mathscr{B}) = \text{ínf } \{d_{\mathscr{B}}(\mathbf{x}) \mid \mathbf{x} \in \mathscr{A}\}.$$

Tendremos ocasión de usar el siguiente

4.1 Teorema. *Si $\mathscr{A}$ es un conjunto cerrado y acotado en R^n y $\mathscr{B}$ es un conjunto cerrado en R^n tal que $\mathscr{A} \cap \mathscr{B} = \varnothing$, entonces $d(\mathscr{A}, \mathscr{B}) > 0$.*

PRUEBA. Como $d_{\mathscr{B}}$ es continua sobre el conjunto cerrado y acotado $\mathscr{A}$, $d_{\mathscr{B}}$ tiene un mínimo sobre $\mathscr{A}$ (teorema 7.9, pág. 479). Es decir, para algún $\mathbf{x}_0 \in \mathscr{A}$, $d(\mathscr{A}, \mathscr{B}) = d_{\mathscr{B}}(\mathbf{x}_0) = d(\mathbf{x}_0, \mathscr{B})$. Como $\mathscr{A} \cap \mathscr{B} = \varnothing$ y $\mathscr{B}$ es cerrado, $d(\mathbf{x}_0, \mathscr{B}) > 0$. Luego $d(\mathscr{A}, \mathscr{B}) > 0$.

4.2 Teorema. (El primer teorema fundamental del cálculo.) *Sea $\mathscr{G}$ un conjunto abierto en R^n y sea f una función de punto continua en $\mathscr{G}$. Si $\mathscr{F}$ es un subconjunto cerrado y acotado de $\mathscr{G}$ que tiene contenido, entonces la función de conjunto F, definida sobre el anillo $\mathfrak{R}$ de subconjuntos de $\mathscr{G}$ que tienen contenido de acuerdo con la regla*

$$F(\mathscr{E}) = \int_{\mathscr{E}} f \qquad \mathscr{E} \in \mathfrak{R}$$

es uniformemente diferenciable sobre $\mathscr{F}$ y $DF = f$ sobre $\mathscr{F}$.

PRUEBA. Sea $\varepsilon > 0$ cualquiera. Deseamos demostrar que existe un número $\delta > 0$ tal que

$$\left| \frac{F(\mathscr{E})}{c(\mathscr{E})} - f(\mathbf{x}) \right| < \varepsilon$$

siempre $\mathbf{x} \in \mathscr{F}$ y $\mathscr{E} \in \mathfrak{R}$ es un conjunto que contiene a $\mathbf{x}$, tiene contenido positivo y diámetro menor que δ.

Como $\mathscr{F}$ es cerrado y acotado y $\mathscr{C}\mathscr{G}$ es cerrado, el teorema 4.1 implica que $d(\mathscr{F}, \mathscr{C}\mathscr{G}) > 0$. Sea $r = \frac{1}{2} d(\mathscr{F}, \mathscr{C}\mathscr{G})$ y sea $\mathscr{H} = \{\mathbf{x} \mid d(\mathbf{x}, \mathscr{F}) \leqslant r\}$. Entonces $\mathscr{H}$ es un conjunto cerrado y acotado y $\mathscr{F} \subset \mathscr{H} \subset \mathscr{G}$. Como f es continua sobre $\mathscr{G}$, f es uniformemente continua sobre $\mathscr{H}$ (teorema 7.6, pág. 478). Luego para cada $\varepsilon > 0$ existe un $\delta_1 > 0$ tal que $|f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon$ siempre que $\mathbf{x}$, $\mathbf{y} \in \mathscr{H}$ y $|\mathbf{x} - \mathbf{y}| < \delta_1$. Sea ahora $\delta = \text{mín } \{\delta_1, r\}$. Entonces $\mathbf{x} \in \mathscr{F}$ y $|\mathbf{x} - \mathbf{y}| < \delta$ implica $\mathbf{y} \in \mathscr{H}$ y $|f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon$.

Tomemos $\mathbf{x} \in \mathscr{F}$ y sea $\mathscr{E} \in \mathfrak{R}$ un conjunto que contiene a $\mathbf{x}$, tiene contenido positivo y diámetro menor que δ. Si $m = \text{ínf } \{f(\mathbf{y}) \mid \mathbf{y} \in \mathscr{E}\}$ y $M = \sup \{f(\mathbf{y}) \mid \mathbf{y} \in \mathscr{E}\}$, entonces

$$0 \leqslant f(\mathbf{x}) - m \leqslant \varepsilon \quad \text{y} \quad 0 \leqslant M - f(\mathbf{x}) \leqslant \varepsilon.$$

Además, por las propiedades 6.2 y 6.7, pág. 341, generalizadas a R^n, tenemos

$$mc(\mathscr{E}) \leqslant F(\mathscr{E}) = \int_{\mathscr{E}} f \leqslant Mc(\mathscr{E}).$$

Por tanto,

$$f(\mathbf{x}) - \varepsilon \leqslant m \leqslant \frac{F(\mathscr{E})}{c(\mathscr{E})} \leqslant M \leqslant f(\mathbf{x}) + \varepsilon$$

o bien

$$\left| \frac{F(\mathscr{E})}{c(\mathscr{E})} - f(\mathbf{x}) \right| \leqslant \varepsilon.$$

Lo que muestra que F es uniformemente diferenciable sobre $\mathscr{F}$ y $DF = f$ sobre $\mathscr{F}$.

Antes de probar el análogo del segundo teorema fundamental del cálculo, probamos dos lemas. El primer lema es un caso especial del *teorema de extensión de Tietze*. Una función g se dice que es una *extensión* de una función f si $\mathscr{D}_f \subset \mathscr{D}_g$ y $f(x) = g(x)$ para todo $x \in \mathscr{D}_f$. Si g es una extensión de f, entonces f es la restricción de g al dominio de f.

4.3 Lema. (Teorema de extensión de Tietze.) *Sea $\mathscr{A}$ un conjunto cerrado en R^n y sea f una función real, acotada y continua definida en $\mathscr{A}$. Entonces existe una función real y continua g definida en R^n que es una extensión de f y es tal que* $\sup\{g(\mathbf{x}) \mid \mathbf{x} \in \mathrm{R}^n\} = \sup\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{A}\} = M$ *e* $\inf\{g(\mathbf{x}) \mid \mathbf{x} \in \mathrm{R}^n\} = \inf\{f(\mathbf{x}) \mid \mathbf{x} \in \mathscr{A}\} = m$.

PRUEBA. Sin pérdida de generalidad podemos suponer que $m > 0$, puesto que en caso contrario podríamos añadir una función constante a f. Definamos g sobre R^n por la regla de correspondencia

$$g(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{para } \mathbf{x} \in \mathscr{A} \\ \dfrac{1}{d(\mathbf{x}, \mathscr{A})} \inf\{f(\mathbf{y})|\mathbf{x}-\mathbf{y}| \quad \mathbf{y} \in \mathscr{A}\} & \text{para } \mathbf{x} \in \mathscr{C}\mathscr{A}. \end{cases}$$

Si $\mathbf{x} \in \mathscr{C}\mathscr{A}$, entonces $d(\mathbf{x}, \mathscr{A}) > 0$, digamos $d(\mathbf{x}, \mathscr{A}) = r$. Para todo $\mathbf{y} \in \mathscr{A}$,

$$m|\mathbf{x}-\mathbf{y}| \leqslant f(\mathbf{y})\,|\mathbf{x}-\mathbf{y}| \leqslant M|\mathbf{x}-\mathbf{y}|.$$

Por tanto

$$\inf_{\mathbf{y} \in \mathscr{A}} \frac{m|\mathbf{x}-\mathbf{y}|}{r} \leqslant \inf_{\mathbf{y} \in \mathscr{A}} \frac{f(\mathbf{y})|\mathbf{x}-\mathbf{y}|}{r} \leqslant \inf_{\mathbf{y} \in \mathscr{A}} \frac{M|\mathbf{x}-\mathbf{y}|}{r}$$

es decir, $m \leqslant g(\mathbf{x}) \leqslant M$ para $\mathbf{x} \in \mathscr{C}\mathscr{A}$. De donde para toda $\mathbf{x} \in \mathrm{R}^n$, $m \leqslant g(\mathbf{x}) \leqslant M$.

Pasamos ahora a probar la continuidad de g. Consideramos puntos de $\mathscr{A}_i$, $\mathscr{A}_e = \mathscr{C}\mathscr{A}$ y $\mathscr{A}_b$ separadamente.

1. La continuidad de g sobre $\mathscr{A}_i$ se sigue inmediatamente de la de f.

2. Sobre el conjunto abierto $\mathscr{C}\mathscr{A} = \mathscr{A}_e$ sea $g(\mathbf{x}) = \dfrac{1}{d(\mathbf{x}, \mathscr{A})} h(\mathbf{x})$ donde $h(\mathbf{x}) = \inf_{\mathbf{y} \in \mathscr{A}} f(\mathbf{y}) |\mathbf{x}-\mathbf{y}|$. Como $d(\mathbf{x}, \mathscr{A}) > 0$ para $\mathbf{x} \in \mathscr{C}\mathscr{A}$ y d es continua sobre $\mathscr{C}\mathscr{A}$, g es continua sobre $\mathscr{C}\mathscr{A}$ si y sólo si h es continua sobre $\mathscr{C}\mathscr{A}$. Sea $d(\mathbf{x}, \mathscr{A}) = r$. Para $\mathbf{x}$, $\mathbf{x}' \in \mathscr{C}\mathscr{A}$ con $|\mathbf{x}-\mathbf{x}'| < \varepsilon \leqslant r$ y para $\mathbf{y} \in \mathscr{A}$, tenemos

$$|\mathbf{x}-\mathbf{y}| < |\mathbf{x}'-\mathbf{y}| + \varepsilon.$$

Luego

$$f(\mathbf{y}) |\mathbf{x}-\mathbf{y}| < f(\mathbf{y}) |\mathbf{x}'-\mathbf{y}| + f(\mathbf{y}) \varepsilon \leqslant f(\mathbf{y}) |\mathbf{x}'-\mathbf{y}| + M\varepsilon$$

y, por tanto,

$$h(\mathbf{x}) \leqslant h(\mathbf{x}') + M\varepsilon.$$

Análogamente, $h(\mathbf{x}') \leqslant h(\mathbf{x}) + M\varepsilon$ de modo que $|h(\mathbf{x}) - h(\mathbf{x}')| \leqslant M\varepsilon$ siempre que $|\mathbf{x}-\mathbf{x}'| < \varepsilon \leqslant r$. Así pues, h es continua sobre $\mathscr{C}\mathscr{A}$ y, por tanto, g es continua sobre $\mathscr{C}\mathscr{A}$.

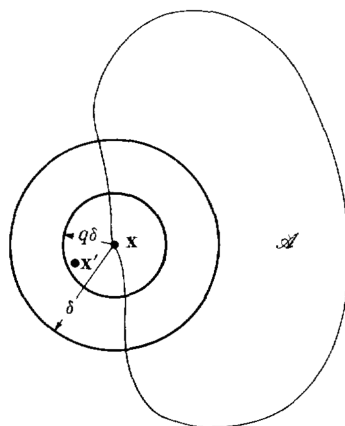


FIGURA 1

3. Para $\mathbf{x} \in \mathscr{A}_b$, dado un $\varepsilon > 0$ sea $\delta > 0$ tal que $\mathbf{y} \in \mathscr{A} \cap \mathscr{S}(\mathbf{x}; \delta)$ implica $|f(\mathbf{y}) - f(\mathbf{x})| < \varepsilon$. Tomemos $\mathbf{x}' \in \mathscr{C}\mathscr{A}$ tal que $|\mathbf{x}-\mathbf{x}'| < q\delta$ donde $q = \dfrac{m}{M+m}$ (figura 1). Deseamos mostrar que

4.4 $$d(\mathbf{x}', \mathscr{A}) = \inf \{|\mathbf{x}'-\mathbf{y}| \mid \mathbf{y} \in \mathscr{A} \cap \mathscr{S}(\mathbf{x}; \delta)\}$$

y

4.5 $$\inf_{\mathbf{y}\in\mathscr{A}} f(\mathbf{y})\,|\mathbf{x}'-\mathbf{y}| = \inf_{\mathbf{y}\in\mathscr{A}\cap\mathscr{S}(\mathbf{x};\,\delta)} f(\mathbf{y})\,|\mathbf{x}'-\mathbf{y}|.$$

Si $\mathbf{y}\in\mathscr{A}-\mathscr{S}(\mathbf{x};\,\delta)$, entonces

$$|\mathbf{x}'-\mathbf{y}| \geqslant |\mathbf{x}-\mathbf{y}| - |\mathbf{x}-\mathbf{x}'| \geqslant (1-q)\delta = \frac{M}{M+m}\,\delta \geqslant q\delta > |\mathbf{x}'-\mathbf{x}|.$$

La ecuación 4.4 se sigue de esta desigualdad. Además

$$\inf_{\mathbf{y}\in\mathscr{A}-\mathscr{S}(\mathbf{x};\,\delta)} f(\mathbf{y})|\mathbf{x}'-\mathbf{y}| \geqslant m(1-q)\delta = Mq\delta \geqslant f(\mathbf{x})|\mathbf{x}'-\mathbf{x}|$$

y la ecuación 4.5 se sigue de esta desigualdad. Ahora bien, para $\mathbf{y}\in\mathscr{A}\cap\mathscr{S}(\mathbf{x};\,\delta)$ y $\mathbf{x}'\in\mathscr{C}\mathscr{A}\cap\mathscr{S}(\mathbf{x};\,q\delta)$, tenemos

$$[f(\mathbf{x})-\varepsilon]\,|\mathbf{x}'-\mathbf{y}| < f(\mathbf{y})\,|\mathbf{x}'-\mathbf{y}| < [f(\mathbf{x})+\varepsilon]\,|\mathbf{x}'-\mathbf{y}|.$$

Por tanto

$$\begin{aligned}\inf_{\mathbf{y}\in\mathscr{A}\cap\mathscr{S}(\mathbf{x};\,\delta)} [f(\mathbf{x})-\varepsilon]\,|\mathbf{x}'-\mathbf{y}| &\leqslant \inf_{\mathbf{y}\in\mathscr{A}\cap\mathscr{S}(\mathbf{x};\,\delta)} f(\mathbf{y})\,|\mathbf{x}'-\mathbf{y}| \\ &\leqslant \inf_{\mathbf{y}\in\mathscr{A}\cap\mathscr{S}(\mathbf{x};\,\delta)} [f(\mathbf{x})+\varepsilon]\,|\mathbf{x}'-\mathbf{y}|\end{aligned}$$

y, por las ecuaciones 4.4 y 4.5,

$$[f(\mathbf{x})-\varepsilon]\,d(\mathbf{x}',\mathscr{A}) \leqslant \inf_{\mathbf{y}\in\mathscr{A}} f(\mathbf{y})\,|\mathbf{x}'-\mathbf{y}| \leqslant [f(\mathbf{x})+\varepsilon]\,d(\mathbf{x}',\mathscr{A})$$

de modo que

$$f(\mathbf{x})-\varepsilon \leqslant g(\mathbf{x}') \leqslant f(\mathbf{x})+\varepsilon.$$

Así pues, para $\mathbf{x}\in\mathscr{C}\mathscr{A}$ y $|\mathbf{x}-\mathbf{x}'| < q\delta$, tenemos

$$|g(\mathbf{x}')-g(\mathbf{x})| = |g(\mathbf{x}')-f(\mathbf{x})| \leqslant \varepsilon.$$

Por otra parte, para $\mathbf{x}'\in\mathscr{A}$ y $|\mathbf{x}-\mathbf{x}'| < q\delta$,

$$|g(\mathbf{x}')-g(\mathbf{x})| = |f(\mathbf{x}')-f(\mathbf{x})| < \varepsilon.$$

Por tanto, $|\mathbf{x}-\mathbf{x}'| < q\delta$ implica $|g(\mathbf{x}')-g(\mathbf{x})| \leqslant \varepsilon$ y g es continua en $\mathbf{x}$. Y esto completa la prueba del lema.

El siguiente lema es un caso particular del segundo teorema fundamental del cálculo.

4.6 Lema. *Sea* $[\mathbf{a},\mathbf{b}]$ *un intervalo en* R^n *y sea* F *una función de conjunto aditiva definida sobre el anillo* $\mathfrak{R}$ *de subconjuntos de* $[\mathbf{a},\mathbf{b}]$ *que tienen contenido. Si* F *es uniformemente diferenciable en* $[\mathbf{a},\mathbf{b}]$ *y si sobre* $[\mathbf{a},\mathbf{b}]$ *la función de*

punto f es igual a DF, entonces f es uniformemente continua sobre $[\mathbf{a}, \mathbf{b}]$ *y para cada intervalo* $\mathcal{J} \subset [\mathbf{a}, \mathbf{b}]$, $F(\mathcal{J}) = \int_{\mathcal{J}} f$.

PRUEBA. Como F es uniformemente diferenciable sobre $[\mathbf{a}, \mathbf{b}]$, para cualquier $\varepsilon > 0$ existe un $\delta > 0$ tal que

$$\left| \frac{F(\mathscr{E})}{c(\mathscr{E})} - f(\mathbf{x}) \right| < \varepsilon$$

siempre que $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$ y $\mathscr{E} \in \mathfrak{R}$ contenga $\mathbf{x}$, tenga contenido positivo y tenga diámetro menor que δ. Sean $\mathbf{x}, \mathbf{y} \in [\mathbf{a}, \mathbf{b}]$ tales que $|\mathbf{x}-\mathbf{y}| < \delta$ y escojamos $\mathscr{E} \in \mathfrak{R}$ de modo tal que $\mathbf{x}, \mathbf{y} \in \mathscr{E}$, $c(\mathscr{E}) > 0$ y $d(\mathscr{E}) < \delta$. Entonces

$$|f(\mathbf{x})-f(\mathbf{y})| \leqslant \left| f(\mathbf{x}) - \frac{F(\mathscr{E})}{c(\mathscr{E})} \right| + \left| \frac{F(\mathscr{E})}{c(\mathscr{E})} - f(\mathbf{y}) \right| < 2\varepsilon$$

y f es uniformemente continua sobre $[\mathbf{a}, \mathbf{b}]$.

Según el teorema de extensión de Tietze, existe una función continua g sobre R^n con valores en R que coincide con f sobre $[\mathbf{a}, \mathbf{b}]$ y es tal que

$$\sup \{g(\mathbf{x}) \mid \mathbf{x} \in \mathrm{R}^n\} = \sup \{f(\mathbf{x}) \mid \mathbf{x} \in [\mathbf{a}, \mathbf{b}]\}$$

e

$$\inf \{g(\mathbf{x}) \mid \mathbf{x} \in \mathrm{R}^n\} = \inf \{f(\mathbf{x}) \mid \mathbf{x} \in [\mathbf{a}, \mathbf{b}]\}.$$

Sea G la función de conjunto definida sobre el anillo $\mathfrak{R}'$ de subconjuntos de R^n que tienen contenido por la regla

$$G(\mathscr{E}) = \int_{\mathscr{E}} g \qquad \mathscr{E} \in \mathfrak{R}'.$$

Según el primer teorema fundamental, G es uniformemente diferenciable sobre $[\mathbf{a}, \mathbf{b}]$ y $DG = g = f$ sobre $[\mathbf{a}, \mathbf{b}]$. Sea $H = F - G$. Sobre $[\mathbf{a}, \mathbf{b}]$, $DH = DF - DG = f - g = 0$ y deseamos demostrar que para cualquier intervalo $\mathcal{J} \subset [\mathbf{a}, \mathbf{b}]$, $F(\mathcal{J}) = G(\mathcal{J})$ o, lo que es equivalente, que $H(\mathcal{J}) = 0$. Tomemos $\mathcal{J} \subset [\mathbf{a}, \mathbf{b}]$. Como $H = F - G$ es uniformemente diferenciable sobre $\mathcal{J}$ con derivada 0, para cada $\varepsilon > 0$ existe una $\delta > 0$ tal que si $\mathscr{E}$ es un subconjunto de $\mathcal{J}$ que tiene contenido distinto de cero y diámetro $d(\mathscr{E}) < \delta$, entonces

$$\left| \frac{H(\mathscr{E})}{c(\mathscr{E})} - 0 \right| < \varepsilon$$

o bien

$$|H(\mathscr{E})| < \varepsilon c(\mathscr{E}).$$

El intervalo $\mathcal{J}$ puede dividirse en un número finito de subintervalos

$\mathscr{I}_j$ $(j = 1, \ldots, k)$ de diámetro menor que δ con $c(\mathscr{I}_i \cap \mathscr{I}_j) = 0$ para $i \neq j$. Ahora bien, G es aditiva, luego $H = F - G$ es aditiva y

$$|H(\mathscr{I})| = \left| \sum_{j=1}^{k} H(\mathscr{I}_j) \right| \leqslant \sum_{j=1}^{k} |H(\mathscr{I}_j)| < \sum_{j=1}^{k} \varepsilon c(\mathscr{I}_j) = \varepsilon c(\mathscr{I}).$$

Como $\varepsilon > 0$ es arbitraria, concluimos que $H(\mathscr{I}) = 0$ para cualquier $\mathscr{I} \subset [\mathbf{a}, \mathbf{b}]$. Por tanto

$$F(\mathscr{I}) = G(\mathscr{I}) = \int_{\mathscr{I}} g = \int_{\mathscr{I}} f.$$

4.7 Teorema. (Segundo teorema fundamental del cálculo.) *Sea* $[\mathbf{a}, \mathbf{b}]$ *un intervalo en* R^n *y sea* F *una función de conjunto monótona y aditiva definida sobre el anillo* $\mathfrak{K}$ *de subconjuntos de* $[\mathbf{a}, \mathbf{b}]$ *que tienen contenido. Si* F *es uniformemente diferenciable sobre* $[\mathbf{a}, \mathbf{b}]$ *y si* $DF = f$ *sobre* $[\mathbf{a}, \mathbf{b}]$, *entonces sobre el anillo*

$$F(\mathscr{E}) = \int_{\mathscr{E}} f \qquad \mathscr{E} \in \mathfrak{K}.$$

PRUEBA. Sea $\mathscr{E} \in \mathfrak{K}$. Como $\mathscr{E}$ tiene contenido, por el teorema 4.17, pág. 337, generalizado a R^n, la frontera $\mathscr{E}_b$ de $\mathscr{E}$ tiene contenido cero. Luego, para un $\varepsilon > 0$ cualquier dado, existe una partición P de $[\mathbf{a}, \mathbf{b}]$ tal que la unión $\mathscr{A}$ de todos aquellos subintervalos de la partición P que contienen puntos de $\mathscr{E}_b$ tiene un contenido $c(\mathscr{A}) < \varepsilon$. Sea $\mathscr{B}$ la unión de todos aquellos subintervalos de la partición P que contienen solamente puntos del interior de $\mathscr{E}$. Como $\mathscr{A}$ y $\mathscr{B}$ son, cada uno, la unión de intervalos y F es aditiva, de acuerdo con el lema 4.6 tenemos

$$F(\mathscr{A}) = \int_{\mathscr{A}} f \quad \text{y} \quad F(\mathscr{B}) = \int_{\mathscr{B}} f.$$

Sea $M = \sup \{|f(\mathbf{x})| \mid \mathbf{x} \in [\mathbf{a}, \mathbf{b}]\}$. Entonces

$$\left| \int_{\mathscr{E}} f - \int_{\mathscr{B}} f \right| = \left| \int_{\mathscr{E} \cap \mathscr{A}} f \right| \leqslant \int_{\mathscr{E} \cap \mathscr{A}} |f| \leqslant Mc(\mathscr{A}) < M\varepsilon$$

y como F es aditiva y monótona,

$$\left| F(\mathscr{E}) - \int_{\mathscr{B}} f \right| = |F(\mathscr{E}) - F(\mathscr{B})| = |F(\mathscr{E} \cap \mathscr{A})| \leqslant |F(\mathscr{A})| = \left| \int_{\mathscr{A}} f \right| < M\varepsilon.$$

Combinando estas dos desigualdades, tenemos

$$\left| F(\mathscr{E}) - \int_{\mathscr{E}} f \right| < 2M\varepsilon.$$

Como $\varepsilon > 0$ es arbitrario, esto nos dice que $F(\mathscr{E}) = \int_{\mathscr{E}} f$.

5. CAMBIO DE VARIABLES EN LAS INTEGRALES MÚLTIPLES. UN CASO ESPECIAL

En el caso de las integrales simples, sabemos que si se efectúa un cambio de variable $y = f(t)$, entonces la integral $\int_a^b g(y)\,dy$ se convierte en la $\int_\alpha^\beta g(f(t))\,f'(t)\,dt$, es decir, tenemos el teorema: si

1) g es continua sobre un intervalo F.
2) f tiene una derivada continua sobre un intervalo $\mathscr{E}$.
3) $f(\mathscr{E}) = \{f(t) \mid t \in \mathscr{E}\} \subset \mathscr{F}$.
4) $f(\alpha) = a$ y $f(\beta) = b$ para algunos α, $\beta \in \mathscr{E}$,

entonces

$$\int_a^b g = \int_\alpha^\beta (g \circ f) f' .$$

En esta sección y en la próxima obtendremos un resultado análogo para las integrales triples.

Ahora daremos una prueba del teorema que para integrales simples acabamos de enunciar. Nuestra prueba del teorema sobre cambio de variable en las integrales triples se modelará sobre esta prueba para integrales simples. Sean

$$G(x) = \int_a^x g \quad \text{y} \quad F(t) = G(f(t)).$$

Entonces, según el primer teorema fundamental del cálculo

$$G'(x) = g(x)$$

y de acuerdo con la regla de la cadena para la derivada de las funciones compuestas,

$$F'(t) = G'(f(t))f'(t) = g(f(t))f'(t) \qquad t \in \mathscr{E}.$$

De donde, según el segundo teorema fundamental del cálculo,

$$\int_\alpha^\beta g(f(t))f'(t)\,dt = \int_\alpha^\beta F'(t)\,dt = F(\beta) - F(\alpha) = G(f(\beta)) - G(f(\alpha))$$

$$= G(b) - G(a) = G(b) = \int_a^b g .$$

En esta prueba hemos usado tanto el primero como el segundo teoremas fundamentales del cálculo y la regla de la cadena. Otro hecho que implícitamente se está usando en la prueba, es el resultado de que la continuidad de f sobre el intervalo $\mathscr{E}$ implica que $f(\mathscr{E})$ es un intervalo. Para la integral

triple consideramos una función de conjunto G definida, sobre un anillo $\mathcal{R}$ de conjuntos que tienen volumen, según la regla de correspondencia

$$G(\mathscr{F}) = \int_{\mathscr{F}} g, \qquad \mathscr{F} \in \mathcal{R}$$

y deseamos obtener un teorema para un cambio de variable dado por una transformación[1] f de R^3 en R^3. En vista de nuestra anunciada intención de dar una prueba modelada sobre la prueba del caso unidimensional, notamos que en la sección 4 obtuvimos los dos teoremas fundamentales del cálculo, el primero y el segundo, para funciones de conjunto. Necesitaremos una regla de la cadena para la derivada de $G \circ \mathbf{f}$, y también será necesario demostrar que las transformaciones que consideramos que transforman conjuntos que tienen volumen en conjuntos que tienen volumen.

En esta sección consideraremos transformaciones lineales de R^3 en R^3. Una transformación $\mathbf{f}$ de R^n en R^m con regla de correspondencia

$$\mathbf{f}(\mathbf{x}) = A\mathbf{x} + \mathbf{y}^0,$$

donde A es una matriz $m \times n$ de constantes y $\mathbf{y}^0$ es un punto en R^m, se llama *transformación lineal.*[2] Si $\mathbf{y}^0 = \mathbf{0}$ la transformación es una *transformación lineal homogénea* y si $\mathbf{y}^0 \neq \mathbf{0}$ la transformación es *no homogénea.* Cada transformación lineal homogénea transforma el origen de R^n en el origen de R^m. Una transformación lineal no homogénea es la composición de una transformación lineal homogénea seguida de una traslación que mueve cada punto de R^m una "cantidad" $\mathbf{y}^0$.

Nota. Es práctica común decir que una transformación $\mathbf{f}$ de R^n en R^m es lineal si

$$\mathbf{f}(r\mathbf{x}+s\mathbf{y}) = r\mathbf{f}(\mathbf{x})+s\mathbf{f}(\mathbf{y})$$

para todo $r, s \in R$ y $\mathbf{x}, \mathbf{y} \in R^n$ y llamar a esta propiedad propiedad de linealidad. Con esta terminología una transformación lineal es homogénea. Una transformación lineal se llama comúnmente *transformación afín.*

Una transformación lineal $\mathbf{f}$ de R^n en R^m es continua sobre R^n. Para cualesquier $\mathbf{x}^1, \mathbf{x}^2 \in R^n$, tenemos

$$|\mathbf{f}(\mathbf{x}^1)-\mathbf{f}(\mathbf{x}^2)| = |A(\mathbf{x}^1-\mathbf{x}^2)| \leqslant \|A\|\,|\mathbf{x}^1-\mathbf{x}^2|$$

donde $\|A\|$ es la norma matricial euclidiana de A, pág. 259. Si $\|A\| = 0$, entonces $\mathbf{f}$ es una transformación constante y $|\mathbf{f}(\mathbf{x}^1)-\mathbf{f}(\mathbf{x}^2)| = 0 < \varepsilon$ para

[1] El término "transformación" es sinónimo del término "función" y a menudo se emplea para funciones de R^n en R^m cuando m y n son, ambos, mayores que uno.

[2] A tal transformación es muy frecuente llamarla *afín*, reservando el nombre de *lineales* a las que los autores llaman *lineales homogéneas.* [N. del T.]

cualesquier $\mathbf{x}^1, \mathbf{x}^2 \in R^n$. Si $\|A\| > 0$, entonces para cada $\varepsilon > 0$ podemos tomar $\delta = \dfrac{\varepsilon}{\|A\|}$. Entonces $|\mathbf{x}^1 - \mathbf{x}^2| < \delta$ implica $|\mathbf{f}(\mathbf{x}^1) - \mathbf{f}(\mathbf{x}^2)| < \varepsilon$.

Sea $\mathbf{f}$ una transformación lineal de R^3 en R^3:

$$\mathbf{f}(\mathbf{x}) = A\mathbf{x} + \mathbf{y}^0$$

donde

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Veamos primero algunos resultados sobre el cambio de volumen de los conjuntos bajo la transformación $\mathbf{f}$.

5.1 Lema. *Sea* $\mathbf{f}$ *una transformación lineal de* R^3 *en* R^3. *Si* $[\mathbf{a}, \mathbf{b}]$ *es un intervalo en* R^3, *entonces* $\mathbf{f}([\mathbf{a}, \mathbf{b}])$ *tiene volumen y*

$$V(\mathbf{f}[\mathbf{a}, \mathbf{b}]) = |\det(A)|\; V([\mathbf{a}, \mathbf{b}])$$

donde $\det(A)$ *denota el determinante de la matriz* A.

PRUEBA. El intervalo $[\mathbf{a}, \mathbf{b}]$ es un paralelepípedo rectangular con lados

$$\mathbf{x}^1 = (b_1 - a_1, 0, 0), \quad \mathbf{x}^2 = (0, b_2 - a_2, 0), \quad \mathbf{x}^3 = (0, 0, b_3 - a_3)$$

paralelos a los ejes coordenados, es decir, $[\mathbf{a}, \mathbf{b}]$ es el conjunto de puntos $\mathbf{x}$ tales que

$$\mathbf{x} = \mathbf{a} + t_1(b_1 - a_1)\mathbf{i} + t_2(b_2 - a_2)\mathbf{j} + t_3(b_3 - a_3)\mathbf{k}$$

donde $t_i \in [0, 1]$, $i = 1, 2, 3$. Como

$$\begin{aligned} \mathbf{f}(\mathbf{x}) = A\mathbf{x} + \mathbf{y}^0 &= A\mathbf{a} + \mathbf{y}^0 + t_1(b_1 - a_1)A\mathbf{i} + t_2(b_2 - a_2)A\mathbf{j} + t_3(b_3 - a_3)A\mathbf{k} \\ &= \mathbf{f}(\mathbf{a}) + t_1(b_1 - a_1)\mathbf{A}^1 + t_2(b_2 - a_2)\mathbf{A}^2 + t_3(b_3 - a_3)\mathbf{A}^3, \end{aligned}$$

$\mathbf{f}([\mathbf{a}, \mathbf{b}])$ es un paralelepípedo de lados

$$\mathbf{y}^i = (b_i - a_i)\mathbf{A}^i, \qquad i = 1, 2, 3$$

donde $\mathbf{A}^i$ es la i-ésima columna de la matriz A. Ahora bien, el volumen del paralelepípedo $\mathbf{f}([\mathbf{a}, \mathbf{b}])$ es el valor absoluto del triple producto escalar de los lados (pág. 61). Por tanto, usando el problema 7b, página 58, obtenemos

$$\begin{aligned} V(\mathbf{f}([\mathbf{a}, \mathbf{b}])) &= |[\mathbf{y}^1\, \mathbf{y}^2\, \mathbf{y}^3]| = |\mathbf{y}^1 \cdot (\mathbf{y}^2 \times \mathbf{y}^3)| \\ &= (b_1 - a_1)(b_2 - a_2)(b_3 - a_3)\, |\mathbf{A}^1 \cdot (\mathbf{A}^2 \times \mathbf{A}^3)| \\ &= V([\mathbf{a}, \mathbf{b}])\, |\det(A^T)| = V([\mathbf{a}, \mathbf{b}])\, |\det(A)| \end{aligned}$$

donde A^T es la transpuesta de la matriz A, es decir, A^T es la matriz que se obtiene escribiendo los renglones de A como las columnas de A^T.

5.2 Teorema. *Sea* f *una transformación de* R^3 *en* R^3 *con regla de correspondencia* $\mathbf{f}(\mathbf{x}) = A\mathbf{x}+\mathbf{y}^0$. *Si* $\mathscr{E} \subset R^3$ *es un conjunto acotado que tiene volumen, entonces* $\mathbf{f}(\mathscr{E})$ *tiene volumen y* $V(\mathbf{f}(\mathscr{E})) = |\det(A)|\, V(\mathscr{E})$.

PRUEBA. Sea $[\mathbf{c}, \mathbf{d}]$ un intervalo en R^3 tal que $\mathscr{E} \subset [\mathbf{c}, \mathbf{d}]$. Como $\mathscr{E}$ tiene volumen, para cada $\varepsilon > 0$ hay una partición P de $[\mathbf{c}, \mathbf{d}]$ tal que

5.3 $$U(1_{\bar{\mathscr{E}}}, P)-\varepsilon < V(\mathscr{E}) < L(1_{\mathscr{E}_i}, P)+\varepsilon.$$

Sea $\mathscr{A}$ la unión de los subintervalos de P que están contenidos en $\mathscr{E}_i$ y sea $\mathscr{B}$ la unión de los subintervalos de P que contienen puntos de $\bar{\mathscr{E}}$. Entonces

$$L(1_{\mathscr{E}_i}, P) = V(\mathscr{A}) \quad \text{y} \quad U(1_{\bar{\mathscr{E}}}, P) = V(\mathscr{B})$$

y la desigualdad 5.3 se hace

5.4 $$V(\mathscr{B})-\varepsilon < V(\mathscr{E}) < V(\mathscr{A})+\varepsilon.$$

Como $\mathscr{A} \subset \mathscr{E} \subset \mathscr{B}$, tenemos $\mathbf{f}(\mathscr{A}) \subset \mathbf{f}(\mathscr{E}) \subset \mathbf{f}(\mathscr{B})$ y por la desigualdad 4.7 y el lema 4.9, pág. 333,

5.5 $$\underline{V}(\mathbf{f}(\mathscr{A})) \leqslant \underline{V}(\mathbf{f}(\mathscr{E})) \leqslant \overline{V}(f(\mathscr{E})) \leqslant \overline{V}(f(\mathscr{B})).$$

Y ahora, por el lema 5.1, como $\mathscr{A}$ y $\mathscr{B}$ son, cada uno, la unión de un conjunto de intervalos que no se traslapan (es decir, $V(\mathscr{R}_i \cap \mathscr{R}_j) = 0$ para $i \neq j$), concluimos que

$$V(\mathbf{f}(\mathscr{A})) = |\det(A)|\, V(\mathscr{A}) \quad \text{y} \quad V(\mathbf{f}(\mathscr{B})) = |\det(A)|\, V(\mathscr{B}).$$

Por tanto, por las desigualdades 5.4 y 5.5,

$$\begin{aligned}\overline{V}(\mathbf{f}(\mathscr{E}))-\underline{V}(\mathbf{f}(\mathscr{E})) &\leqslant V(\mathbf{f}(\mathscr{B}))-V(\mathbf{f}(\mathscr{A}))\\ &= |\det(A)|\,[V(\mathscr{B})-V(\mathscr{A})] < 2\varepsilon\, |\det(A)|\end{aligned}$$

y como $\varepsilon > 0$ es arbitrario, $\overline{V}(\mathbf{f}(\mathscr{E})) = \underline{V}(\mathbf{f}(\mathscr{E}))$ y $\mathbf{f}(\mathscr{E})$ tiene volumen. De donde la desigualdad 5.5 toma la forma

$$|\det(A)|\, V(\mathscr{A}) \leqslant V(\mathbf{f}(\mathscr{E})) \leqslant |\det(A)|\, V(\mathscr{B}).$$

Como $\mathscr{A} \subset \mathscr{E} \subset \mathscr{B}$ y todas tienen volumen, se tiene

$$|\det(A)|\, V(\mathscr{A}) \leqslant |\det(A)|\, V(\mathscr{E}) \leqslant |\det(A)|\, V(\mathscr{B}).$$

De estas dos últimas desigualdades y de 5.4, obtenemos

$$|V(\mathbf{f}(\mathscr{E}))-|\det(A)|\, V(\mathscr{E})| \leqslant |\det(A)|\,[V(\mathscr{B})-V(\mathscr{A})] < 2\varepsilon\, |\det(A)|$$

y de nuevo, como $\varepsilon > 0$ es arbitraria, concluimos

$$V(\mathbf{f}(\mathscr{E})) = |\det(A)|\, V(\mathscr{E}).$$

5.6 Ejemplo. Sea $\mathscr{F}$ la región limitada por el plano con ecuación $z = x+2y+1$ por la parte superior, debajo por el plano XY, y lateralmente por el cilindro elíptico de ecuación $2x^3+4xy+5y^2+2x-y-4 = 0$. Encuéntrese el volumen de $\mathscr{F}$.

SOLUCIÓN. Buscaremos una transformación $\mathbf{f}$ con regla de correspondencia de la forma $\mathbf{y} = \mathbf{f}(\mathbf{x}) = A\mathbf{x}+\mathbf{y}^0$ tal que $\mathscr{F} = \mathbf{f}(\mathscr{E})$ con una $\mathscr{E}$ que tenga una descripción relativamente simple. Como la frontera lateral de $\mathscr{F}$ es un cilindro elíptico, hay cilindros elípticos de la misma forma con ejes sobre el eje Z y los ejes de una sección normal paralelos a los ejes de coordenadas. Luego podemos considerar a $\mathscr{F}$ como la imagen de una región $\mathscr{E}$ con frontera lateral del tipo descrito al aplicársele una rotación alrededor del eje Z seguida de una traslación. Bajo la transformación con regla de correspondencia (véase el vol. I, pág. 257)

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\dfrac{2}{\sqrt{5}} & -\dfrac{1}{\sqrt{5}} & 0 \\ \dfrac{1}{\sqrt{5}} & -\dfrac{2}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} + \begin{pmatrix} -1 \\ \frac{1}{2} \\ 0 \end{pmatrix},$$

encontramos que la frontera lateral de $\mathscr{F}$ es la imagen del cilindro elíptico con ecuación $u^2+6v^2 = \frac{21}{4}$. La descripción de $\mathscr{E}$ se simplifica aún más si la frontera lateral es un cilindro circular. Reemplazando v en la transformación anterior por $\dfrac{1}{\sqrt{6}}v$, es fácil verificar que la transformación $\mathbf{f}$ con regla de correspondencia

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{f}(u, v, w) = A \begin{pmatrix} u \\ v \\ w \end{pmatrix} + \begin{pmatrix} -1 \\ \frac{1}{2} \\ 0 \end{pmatrix}$$

donde

$$A = \begin{pmatrix} -\dfrac{2}{\sqrt{5}} & -\dfrac{1}{\sqrt{30}} & 0 \\ \dfrac{1}{\sqrt{5}} & -\dfrac{2}{\sqrt{30}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

transforma la región $\mathscr{E}$, limitada lateralmente por el cilindro circular de ecuación $u^2+v^2 = \frac{21}{4}$, por arriba por el plano de ecuación $w = -\frac{1}{6}\sqrt{30}v+1$, y por debajo por el plano UV, en la región $\mathscr{F}$. Luego, por el teorema 5.2,

$$V(\mathscr{F}) = |\det(A)|\, V(\mathscr{E}) = \frac{1}{\sqrt{6}}\, V(\mathscr{E})$$

$$= \frac{1}{\sqrt{6}} \int_{-\frac{1}{2}\sqrt{21}}^{\frac{1}{2}\sqrt{21}} \int_{-\sqrt{21/4-u^2}}^{\sqrt{21/4-u^2}} \int_0^{-\frac{1}{6}\sqrt{30}v+1} dw\, dv\, du = \tfrac{7}{8}\sqrt{6}\,\pi .$$

Consideraremos a continuación una regla de la cadena. Una función de conjunto G definida sobre una familia $\mathfrak{R}$ de conjuntos en R^3 que tienen volumen, es diferenciable sobre un conjunto $\mathscr{S} \subset R^3$ si para cada $\mathbf{y} \in \mathscr{S}$,

$$DG(\mathbf{y}) = \lim_{\mathscr{F}\to\mathbf{y}} \frac{G(\mathscr{F})}{V(\mathscr{F})}$$

existe. Luego si G es diferenciable sobre $\mathscr{S}$, entonces

5.7 $$G(\mathscr{F}) = DG(\mathbf{y})\, V(\mathscr{F}) + \Phi(\mathbf{y}; \mathscr{F})\, V(\mathscr{F})$$

donde $DG(\mathbf{y})$ es independiente de $\mathscr{F}$, la función Φ está definida sobre $\mathscr{S} \times \mathfrak{R}$, y $\lim_{\mathscr{F}\to\mathbf{y}} \Phi(\mathbf{y}; \mathscr{F}) = 0$ para todo $\mathbf{y} \in \mathscr{S}$. Sea $\mathbf{f}$ una transformación lineal de R^3 en R^3. Si $\mathscr{E}$ es un conjunto acotado en R^3 que tiene volumen, entonces, por el teorema 5.2, $\mathscr{F} = \mathbf{f}(\mathscr{E})$ tiene volumen. Definamos la función de conjunto $F = G \circ \mathbf{f}$ por la regla de correspondencia

$$F(\mathscr{E}) = G(\mathbf{f}(\mathscr{E}))$$

donde $\mathscr{E}$ es un conjunto acotado en R^3 que tiene volumen y $\mathbf{f}(\mathscr{E}) \in \mathfrak{D}_G$. Supongamos que $\mathbf{x} \in R^3$ es un punto tal que $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in \mathscr{S}$. Entonces, como $V(\mathscr{F}) = |\det(A)|\, V(\mathscr{E})$, la ecuación 5.7 toma la forma

5.8 $$F(\mathscr{E}) = G(\mathbf{f}(\mathscr{E})) = DG(\mathbf{f}(\mathbf{x}))\, |\det(A)|\, V(\mathscr{E}) + \Phi(\mathbf{f}(\mathbf{x}); \mathbf{f}(\mathscr{E}))\, |\det(A)|\, V(\mathscr{E}).$$

Por otra parte F es diferenciable en $\mathbf{x}$ si y sólo si

$$F(\mathscr{E}) = DF(\mathbf{x})\, V(\mathscr{E}) + \Psi(\mathbf{x}; \mathscr{E})\, V(\mathscr{E})$$

donde $DF(\mathbf{x})$ es independiente de $\mathscr{E}$ y $\lim_{\mathscr{E}\to\mathbf{x}} \Psi(\mathbf{x}; \mathscr{E}) = 0$. Comparando esto con la ecuación 5.8 vemos que F es diferenciable en $\mathbf{x}$ con

5.9 $$DF(\mathbf{x}) = DG(\mathbf{f}(\mathbf{x}))\, |\det(A)|$$

y $\Psi(\mathbf{x}; \mathscr{E}) = \Phi(\mathbf{f}(\mathbf{x}); \mathbf{f}(\mathscr{E}))\, |\det(A)|$. La ecuación 5.9 es la regla de la cadena que buscábamos.

Recordemos ahora la definición de función univalente o uno-uno.[1] Una función es un conjunto de pares ordenados tal que no hay dos pares distintos que tengan el mismo primer elemento. Si, además, no hay pares distintos que tengan el mismo segundo elemento, la función se dice que es *univalente* o *uno-uno*. Una función univalente establece una correspondencia uno-uno entre su dominio y su rango.

5.10 Lema. *Sea* $\mathbf{f}$ *una transformación lineal de* R^3 *en* R^3 *con regla de correspondencia* $\mathbf{f}(\mathbf{x}) = A\mathbf{x}+\mathbf{y}^0$. *Si* $\det(A) \neq 0$, *entonces* $\mathbf{f}$ *es una transformación univalente sobre* R^3 *y si* $\det(A) = 0$, *entonces* $\mathbf{f}$ *transforma a todos los puntos de* R^3 *sobre un plano que pasa por* $\mathbf{y}^0$, *sobre una recta que pasa por* $\mathbf{y}^0$ *o sobre* $\mathbf{y}^0$.

PRUEBA. Sea $\mathbf{y} = \mathbf{f}(\mathbf{x})$, entonces

$$A\mathbf{x} = \mathbf{y}-\mathbf{y}^0.$$

Si $\det(A) \neq 0$, por el corolario 10.4, pág. 80, este sistema de tres ecuaciones lineales tiene una solución única $\mathbf{x} = \mathbf{f}^*(\mathbf{y})$ para cada $\mathbf{y}\in\mathrm{R}^3$. Luego $\mathbf{f}$ es univalente y sobre R^3 (es decir, el rango de $\mathbf{f}$ es todo R^3). Si el $\det(A) = 0$ y $\mathbf{A}^1$, $\mathbf{A}^2$, $\mathbf{A}^3$ son las columnas de $\mathbf{A}$, entonces, por el teorema 6.4, pág. 64, $\mathbf{A}^1$, $\mathbf{A}^2$, $\mathbf{A}^3$ son linealmente dependientes. Ahora bien

$$\mathbf{f}(\mathbf{x}) = \mathbf{y}^0+A\mathbf{x} = \mathbf{y}^0+\mathbf{A}^1x_1+\mathbf{A}^2x_2+\mathbf{A}^3x_3.$$

Si dos de los vectores, digamos $\mathbf{A}^1$, $\mathbf{A}^2$, son linealmente independientes mientras que $\mathbf{A}^1$, $\mathbf{A}^2$, $\mathbf{A}^3$ son linealmente dependientes, según el problema 4, pág. 66, hay número s, $t\in\mathrm{R}$ tales que $\mathbf{A}^3 = s\mathbf{A}^1+t\mathbf{A}^2$ y $\mathbf{f}(\mathbf{x})$ pertenece al plano que pasa por $\mathbf{y}^0$ paralelo a los vectores $\mathbf{A}^1$ y $\mathbf{A}^2$. Si en $\mathbf{A}^1$, $\mathbf{A}^2$, $\mathbf{A}^3$ no hay ningún par de independientes, pero uno de ellos, digamos $\mathbf{A}^1$, es distinto de 0, entonces hay números s, $t\in\mathrm{R}$ tales que $\mathbf{A}^2 = s\mathbf{A}^1$ y $\mathbf{A}^3 = t\mathbf{A}^1$. Entonces $\mathbf{f}(\mathbf{x})$ pertenece a la recta que pasa por $\mathbf{y}^0$ paralela a $\mathbf{A}^1$. Si $\mathbf{A}^1 = \mathbf{A}^2 = \mathbf{A}^3 = \mathbf{0}$, entonces $\mathbf{f}(\mathbf{x}) = \mathbf{y}^0$ para todo $\mathbf{x}\in\mathrm{R}^3$.

Estamos ya en posibilidad de probar un teorema sobre cambio de variable en las integrales triples para cambio lineal de variable.

5.11 Teorema. *Sea* $\mathbf{f}$ *una transformación lineal de* R^3 *en* R^3 *con regla de correspondencia* $\mathbf{f}(\mathbf{x}) = A\mathbf{x}+\mathbf{y}^0$. *Sea* $[\mathbf{a}, \mathbf{b}]$ *un intervalo en* R^3 *y sea* g *una función de* R^3 *en* R *continua sobre un conjunto abierto que contiene a* $\mathbf{f}([\mathbf{a}, \mathbf{b}])$. *Si* $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$ *tiene volumen y* $\mathscr{F} = \mathbf{f}(\mathscr{E})$, *entonces*

$$\int_{\mathscr{F}} g = \int_{\mathscr{F}} g(\mathbf{y})\,d\mathbf{y} = \int_{\mathscr{E}} (g\circ\mathbf{f})\,|\det(A)| = \int_{\mathscr{E}} g(\mathbf{f}(\mathbf{x}))\,|\det(A)|\,d\mathbf{x}.$$

[1] Llamadas también "inyectivas". [N. del T.]

PRUEBA. Si $\det(A) = 0$, el teorema es cierto pero trivial ya que $\mathscr{F}$ se encuentra sobre un plano, recta o punto y $V(\mathscr{F}) = 0$. Supondremos, pues, que $\det(A) \neq 0$. Definamos la función de conjunto G, sobre el anillo $\mathfrak{R}$ de subconjuntos de $\mathbf{f}([\mathbf{a}, \mathbf{b}])$ que tienen volumen, por la regla de correspondencia

$$G(\mathscr{F}) = \int_{\mathscr{F}} g,$$

y sea $F(\mathscr{E}) = G(\mathbf{f}(\mathscr{E}))$. Si g es una función constante, el teorema se sigue directamente del teorema 5.2. Como G es continua sobre el conjunto cerrado y acotado $\mathbf{f}([\mathbf{a}, \mathbf{b}])$, por el teorema 7.7, pág. 478, g es acotada en tal conjunto. Podemos suponer que g sólo toma valores positivos sobre $\mathbf{f}([\mathbf{a}, \mathbf{b}])$ pues en caso contrario $g-m+1$ donde $m = \text{mín}\{g(\mathbf{y}) \mid \mathbf{y}\in\mathbf{f}([\mathbf{a}, \mathbf{b}])\}$. Luego, después de demostrar que el teorema se verifica para la función con todos sus valores positivos $g-m+1$ y para la función constante $m-1$, obtendríamos por adición el teorema para $g = (g-m+1)+(m-1)$. Por el primer teorema fundamental del cálculo, $DG(\mathbf{y}) = g(\mathbf{y})$ para todo $\mathbf{y}\in\mathbf{f}([\mathbf{a}, \mathbf{b}])$ y, por la regla de la cadena (ecuación 5.9), F es diferenciable sobre $[\mathbf{a}, \mathbf{b}]$ con

$$DF(\mathbf{x}) = DG(\mathbf{f}(\mathbf{x}))\,|\det(A)| = g(\mathbf{f}(\mathbf{x}))\,|\det(A)|.$$

Demostraremos que F satisface las condiciones del segundo teorema fundamental del cálculo. El teorema se sigue entonces del segundo teorema fundamental del cálculo. Como el $\det(A) \neq 0$, $\mathbf{f}$ es univalente y $V(\mathscr{E}_1 \cap \mathscr{E}_2) = 0$ implica que $V(\mathbf{f}(\mathscr{E}_1) \cap \mathbf{f}(\mathscr{E}_2)) = V(\mathbf{f}(\mathscr{E}_1 \cap \mathscr{E}_2)) = 0$ (teorema 5.2). Luego F es aditiva ya que lo es G. Como estamos suponiendo que g toma sólo valores positivos, F es monótona. Por el primer teorema fundamental, G es uniformemente diferenciable sobre $\mathbf{f}([\mathbf{a}, \mathbf{b}])$ y, por tanto, la ecuación 5.8 implica que F es uniformemente diferenciable sobre $[\mathbf{a}, \mathbf{b}]$. De donde F satisface las condiciones del segundo teorema fundamental del cálculo e

$$\int_{\mathscr{F}} g = G(\mathscr{F}) = G(\mathbf{f}(\mathscr{E})) = F(\mathscr{E}) = \int_{\mathscr{E}} DF(\mathbf{x})\,d\mathbf{x} = \int_{\mathscr{E}} g(\mathbf{f}(\mathbf{x}))\,|\det(A)|\,d\mathbf{x}.$$

5.12 Ejemplo. Encuéntrese el momento de inercia con respecto al eje Z de la región $\mathscr{F}$ limitada superiormente por el plano de ecuación $z = x+2y+1$, inferiormente por el plano XY, y lateralmente por el cilindro elíptico de ecuación

$$2x^2+4xy+5y^2+2x-y-4 = 0.$$

SOLUCIÓN. La región $\mathscr{F}$ es la misma que la del ejemplo 5.6. Si

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{f}(u, v, w) = A\begin{pmatrix} u \\ v \\ w \end{pmatrix} + \begin{pmatrix} -1 \\ \frac{1}{2} \\ 0 \end{pmatrix}$$

donde

$$A = \begin{pmatrix} -\dfrac{2}{\sqrt{5}} & -\dfrac{1}{\sqrt{30}} & 0 \\ \dfrac{1}{\sqrt{5}} & -\dfrac{2}{\sqrt{30}} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

entonces $\mathscr{F}$ es la imagen bajo $\mathbf{f}$ de la región $\mathscr{E}$ acotada superiormente por el plano de ecuación $w = -\frac{1}{6}\sqrt{30}v+1$, debajo por el plano UV, y lateralmente por el cilindro circular de ecuación $u^2+v^2 = \frac{21}{4}$. El momento de inercia de $\mathscr{F}$ con respecto al eje Z es $\int_{\mathscr{F}} g$ donde $g(x,y,z) = x^2+y^2$. Luego

$$g(\mathbf{f}(u,v,w)) = \left(-\frac{2}{\sqrt{5}}u - \frac{1}{\sqrt{30}}v-1\right)^2 + \left(\frac{1}{\sqrt{5}}u - \frac{2}{\sqrt{30}}v + \frac{1}{2}\right)^2$$

$$= u^2+6v^2+\sqrt{5}u+\tfrac{5}{4}$$

y según el teorema 5.11, el momento de inercia de $\mathscr{F}$ con respecto al eje Z es

$$\int_{\mathscr{F}} g = \int_{\mathbf{f}(\mathscr{E})} g = \int_{\mathscr{E}} g(\mathbf{f}(u,v,w))\,|\det(A)|\,dw\,dv\,du$$

$$= \frac{1}{\sqrt{6}}\int_{-(1/2)\sqrt{21}}^{(1/2)\sqrt{21}}\int_{-\sqrt{21/4-u^2}}^{\sqrt{21/4-u^2}}\int_0^{-(1/6)\sqrt{30}v+1} [u^2+\tfrac{1}{6}v^2+\sqrt{5}u+\tfrac{5}{4}]\,dw\,dv\,du$$

$$= \tfrac{623}{256}\sqrt{6}\pi\,.$$

Problemas

1. Un conjunto de puntos de la forma

$$\mathscr{P} = \{\mathbf{P}_0+u\mathbf{a}+v\mathbf{b}+w\mathbf{c} \mid u,\, v,\, w\in[0,\, 1]\}$$

es un paralelepípedo con un vértice en $\mathbf{P}_0$ y lados $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$. La transformación

$$\mathbf{P} = \mathbf{f}(u,v,w) = \mathbf{P}_0+u\mathbf{a}+v\mathbf{b}+w\mathbf{c} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\begin{pmatrix} u \\ v \\ w \end{pmatrix} + \mathbf{P}_0$$

transforma el cubo unitario $[(0,0,0),\,(1,1,1)]$ sobre $\mathscr{P}$.

a) Usando la transformación **f**, encuéntrese el volumen de $\mathscr{P}$.
b) Encuéntrese el momento de $\mathscr{P}$ con respecto al plano XY.
c) Encuéntrese el momento de inercia de $\mathscr{P}$ con respecto al plano XY.

2. Encuéntrese la transformación lineal **f** que transforma la esfera unitaria con centro en el origen en el elipsoide de ecuación $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1$ y encuéntrese el volumen del elipsoide usando la bien conocida fórmula para el volumen de la esfera.

3. La transformación **f** tal que

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{f}(u,v,w) = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$$

transforma una región $\mathscr{E}$ en una región $\mathscr{F}$ limitada inferiormente por el plano XY, lateralmente por el cilindro parabólico de ecuación

$$x^2-2xy+y^2+x-3y+6 = 0$$

y el plano de ecuación $x-3y+15=0$, y superiormente por el paraboloide de revolución de ecuación $z=x^2+y^2$. Encuéntrese la región $\mathscr{E}$ y úsese la transformación **f** para encontrar el volumen de $\mathscr{F}$.

6. CAMBIO DE VARIABLE EN UNA INTEGRAL MÚLTIPLE

En la sección previa consideramos el teorema para el cambio de variable en integrales triples para cambio lineal de variable. En esta sección consideraremos este mismo problema para un tipo más general de transformación de R^3 en R^3.

Si una transformación **f** es diferenciable en $\mathbf{x}^0$, entonces **f** está definida en alguna vecindad de $\mathbf{x}^0$ y para **x** en esta vecindad

$$\begin{aligned}\mathbf{f}(\mathbf{x}) &= \mathbf{f}(\mathbf{x}^0)+A(\mathbf{x}-\mathbf{x}^0)+\Phi(\mathbf{x}^0;\mathbf{x}-\mathbf{x}^0)(\mathbf{x}-\mathbf{x}^0)\\ &= A\mathbf{x}+[\mathbf{f}(\mathbf{x}^0)-A\mathbf{x}^0]+\Phi(\mathbf{x}^0;\mathbf{x}-\mathbf{x}^0)(\mathbf{x}-\mathbf{x}^0)\end{aligned}$$

donde $\lim\limits_{\mathbf{x}\to\mathbf{x}^0}\Phi(\mathbf{x}^0;\mathbf{x}-\mathbf{x}^0)=0$ y $A=D\mathbf{f}(\mathbf{x}^0)$ es independiente de **x**. Si **f** es una transformación lineal, $\mathbf{y}=\mathbf{f}(\mathbf{x})=A\mathbf{x}+\mathbf{y}^0$, donde A es una función matricial de valores constantes, entonces $D\mathbf{f}(\mathbf{x})=A$ y $\Phi=0$. Sabemos pue bajo tal transformación el volumen está multiplicado por el factor $|\det(A)|$ y que

$$\int_{\mathbf{f}(\mathscr{E})} g = \int_{\mathscr{E}} (g\circ\mathbf{f})\,|\det(A)|\,.$$

En el caso de una función $\mathbf{f}$ diferenciable en $\mathbf{x}^0$, tenemos

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) \approx [D\mathbf{f}(\mathbf{x}^0)]\mathbf{x} + \mathbf{y}^0$$

para $|\mathbf{x}-\mathbf{x}^0|$ suficientemente pequeño donde $\mathbf{y}^0 = \mathbf{f}(\mathbf{x}^0) - [D\mathbf{f}(\mathbf{x}^0)]\mathbf{x}^0$. De todo esto resulta razonable esperar que $|\det(D\mathbf{f}(\mathbf{x}^0))|$ debe reemplazar a $|\det(A)|$ en la fórmula para el cambio de volumen, al menos localmente, y en la fórmula para el cambio de variable en la integral. Que esta conjetura es correcta cuando la transformación $\mathbf{f}$ cumple determinadas condiciones es lo que demostraremos en esta sección.

Nota. El determinante de la derivada se llama ***determinante jacobiano*** de $\mathbf{f}$ en $\mathbf{x}$ o, simplemente, ***jacobiano*** de $\mathbf{f}$ en $\mathbf{x}$ y se denota por $J\mathbf{f}(\mathbf{x})$. Si $(x, y, z) = \mathbf{f}(u, v, w)$, entonces el jacobiano se denota tambien por $\dfrac{\partial(x,y,z)}{\partial(u,v,w)}$.

En la prueba del teorema sobre cambio de variable en las integrales triples para cambio lineal de variable, usamos una regla de la cadena. Esta regla de la cadena se obtiene al expresar el volumen de un conjunto imagen $\mathscr{F} = \mathbf{f}(\mathscr{E})$ en términos del volumen del conjunto $\mathscr{E}$. Para llevar al cabo la prueba de la fórmula del cambio de variable para una transformación más general $\mathbf{f}$ investigaremos la relación del volumen de $\mathscr{F} = \mathbf{f}(\mathscr{E})$ con el volumen de $\mathscr{E}$. Estableceremos primero algunos resultados preliminares.

6.1 Lema. *Si una transformación* $\mathbf{f}$ *de* R^3 *en* R^3 *es de clase* C^1 *sobre un conjunto abierto* $\mathscr{G}$ *y el jacobiano* $J\mathbf{f}(\mathbf{x}) \neq 0$ *para todo* $\mathbf{x} \in \mathscr{G}$, *entonces* $\mathbf{f}$ *es localmente univalente, es decir, para cada* $\mathbf{x} \in \mathscr{G}$ *hay una vecindad* $\mathscr{S}(\mathbf{x}; \delta) \subset \mathscr{G}$ *tal que* $\mathbf{f}$ *es univalente sobre* $\mathscr{S}(\mathbf{x}; \delta)$.

PRUEBA. Sea $\mathbf{x}$ un punto de $\mathscr{G}$ y sea $\mathscr{S}(\mathbf{x}) \subset \mathscr{G}$ una vecindad de $\mathbf{x}$. Si $\mathbf{y}^1$, $\mathbf{y}^2 \in \mathscr{S}(\mathbf{x})$, entonces el segmento que une $\mathbf{y}^1$ y $\mathbf{y}^2$ pertenece a $\mathscr{S}(\mathbf{x}) \in \mathscr{G}$. Por el teorema del valor medio (teorema 6.10, pág. 194), existe un punto $\mathbf{z}^i \in \langle \mathbf{y}^1, \mathbf{y}^2 \rangle$ tal que

6.2
$$\mathbf{f}(\mathbf{y}^2) - \mathbf{f}(\mathbf{y}^1) = \begin{pmatrix} D_1 f_1(\mathbf{z}^1) & D_2 f_1(\mathbf{z}^1) & D_3 f_1(\mathbf{z}^1) \\ D_1 f_2(\mathbf{z}^2) & D_2 f_2(\mathbf{z}^2) & D_3 f_2(\mathbf{z}^2) \\ D_1 f_3(\mathbf{z}^3) & D_2 f_3(\mathbf{z}^3) & D_3 f_3(\mathbf{z}^3) \end{pmatrix} (\mathbf{y}^2 - \mathbf{y}^1)$$

$$= A(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^3)(\mathbf{y}^2 - \mathbf{y}^1)$$

$$= \begin{pmatrix} \mathbf{D}f_1(\mathbf{z}^1) \cdot (\mathbf{y}^2 - \mathbf{y}^1) \\ \mathbf{D}f_2(\mathbf{z}^2) \cdot (\mathbf{y}^2 - \mathbf{y}^1) \\ \mathbf{D}f_3(\mathbf{z}^3) \cdot (\mathbf{y}^2 - \mathbf{y}^1) \end{pmatrix}.$$

Como $\det(A(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^3)) = \mathbf{D}f_1(\mathbf{z}^1) \cdot \mathbf{D}f_2(\mathbf{z}^2) \times \mathbf{D}f_3(\mathbf{z}^3)$ es una suma de productos de $D_j f_i(\mathbf{z}^i)$ y las derivadas parciales $D_j f_i$ son continuas sobre $\mathscr{G}$, $\det(A)$ es continua sobre $\mathscr{S}(\mathbf{x}) \times \mathscr{S}(\mathbf{x}) \times \mathscr{S}(\mathbf{x})$. Como $\det(A(\mathbf{x}, \mathbf{x}, \mathbf{x})) = J\mathbf{f}(\mathbf{x}) \neq 0$ hay un $\delta > 0$ tal que si $\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^3 \in \mathscr{S}(\mathbf{x}; \delta) \subset \mathscr{S}(\mathbf{x})$ entonces $\det(A(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^3)) \neq 0$. Tomemos $\mathbf{y}^1, \mathbf{y}^2 \in \mathscr{S}(\mathbf{x}; \delta)$. Entonces, en la ecuación 6.2, $\det(A(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^3)) = \mathbf{D}f_1(\mathbf{z}^1) \cdot \mathbf{D}f_2(\mathbf{z}^2) \times \mathbf{D}f_3(\mathbf{z}^3) \neq 0$.

Si $\mathbf{f}(\mathbf{y}^2) - \mathbf{f}(\mathbf{y}^1) = \mathbf{0}$, entonces, por la ecuación 6.2, vemos que

$$A(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^3)(\mathbf{y}^2 - \mathbf{y}^1) = \mathbf{0}.$$

Pero como el determinante de los coeficientes en este sistema de ecuaciones lineales es distinto de cero, $\mathbf{y}^2 - \mathbf{y}^1 = \mathbf{0}$ (corolario 10.4, pág. 80). Así pues, $\mathbf{y}^1, \mathbf{y}^2 \in \mathscr{S}(\mathbf{x}; \delta)$ y $\mathbf{f}(\mathbf{y}^2) = \mathbf{f}(\mathbf{y}^1)$ implican $\mathbf{y}^2 = \mathbf{y}^1$ y esto demuestra que $\mathbf{f}$ es univalente sobre $\mathscr{S}(\mathbf{x}; \delta)$.

6.3 Teorema. *Si una transformación* $\mathbf{f}$ *de* R^3 *en* R^3 *es de clase* C^1 *sobre un conjunto abierto* $\mathscr{G}$ *y el jacobiano* $J\mathbf{f}(\mathbf{x}) \neq 0$ *para todo* $\mathbf{x} \in \mathscr{G}$, *entonces* $\mathbf{f}(\mathscr{G})$ *es abierto.*

PRUEBA. Tomemos $\mathbf{y}^0 \in \mathbf{f}(\mathscr{G})$. Deseamos demostrar que existe una vecindad de $\mathbf{y}^0$ que está contenida en $\mathbf{f}(\mathscr{G})$. Sea $\mathbf{x}^0 \in \mathscr{G}$ un punto tal que $\mathbf{f}(\mathbf{x}^0) = \mathbf{y}^0$. Por el lema 6.1, existe una vecindad $\mathscr{S}$ de $\mathbf{x}^0$ cuya cerradura $\overline{\mathscr{S}}$ está contenida en $\mathscr{G}$ tal que $\mathbf{f}$ es univalente sobre $\overline{\mathscr{S}}$. Por el teorema 7.8, pág. 478, $\mathbf{f}(\mathscr{S}_b)$ es un conjunto cerrado y acotado. Sea $d = d(\mathbf{y}^0, \mathbf{f}(\mathscr{S}_b))$. Como $\mathbf{y}^0 \notin \mathbf{f}(\mathscr{S}_b)$ y $\mathbf{f}(\mathscr{S}_b)$ es cerrado, $d > 0$. Demostraremos que si $|\mathbf{y}^1 - \mathbf{y}^0| < d/2$, entonces $\mathbf{y}^1 \in \mathbf{f}(\mathscr{S}) \subset \mathbf{f}(\mathscr{G})$. Es decir, probaremos que existe un $\mathbf{x}^1 \in \mathscr{S}$ tal que $\mathbf{f}(\mathbf{x}^1) = \mathbf{y}^1$. Sea g la función definida sobre $\overline{\mathscr{S}}$ por la regla

$$g(\mathbf{x}) = |\mathbf{f}(\mathbf{x}) - \mathbf{y}^1|^2 = (\mathbf{f}(\mathbf{x}) - \mathbf{y}^1) \cdot (\mathbf{f}(\mathbf{x}) - \mathbf{y}^1);$$

$g(\mathbf{x})$ es el cuadrado de la distancia de $\mathbf{f}(\mathbf{x})$ a $\mathbf{y}^1$. Deseamos probar que hay un $\mathbf{x}^1 \in \mathscr{S}$ tal que $g(\mathbf{x}^1) = 0$, es decir, tal que $\mathbf{f}(\mathbf{x}^1) = \mathbf{y}^1$. Como g es continua sobre el conjunto cerrado y acotado $\overline{\mathscr{S}}$, por el teorema 7.9, pág. 479, g tiene un valor mínimo sobre $\overline{\mathscr{S}}$. Por otra parte, este mínimo no ocurre en un punto de $\mathscr{S}_b$ ya que si $\mathbf{x} \in \mathscr{S}_b$, tenemos

$$|\mathbf{f}(\mathbf{x}) - \mathbf{y}^1| \geqslant |\mathbf{f}(\mathbf{x}) - \mathbf{y}^0| - |\mathbf{y}^0 - \mathbf{y}^1| > d - \tfrac{1}{2}d = \tfrac{1}{2}d$$

mientras que

$$|\mathbf{f}(\mathbf{x}^0) - \mathbf{y}^1| = |\mathbf{y}^0 - \mathbf{y}^1| < \tfrac{1}{2}d.$$

Como $\mathbf{f}$ es de clase C^1 sobre $\mathscr{G}$, g es de clase C^1 sobre $\mathscr{S}$ y el mínimo de g se presenta en un punto $\mathbf{x}^1$ donde todas las derivadas parciales de g se anulan:

$$D_j g(\mathbf{x}^1) = 2(\mathbf{f}(\mathbf{x}^1) - \mathbf{y}^1) \cdot D_j \mathbf{f}(\mathbf{x}^1) = 0. \quad (j = 1, 2, 3).$$

Es este un sistema de tres ecuaciones lineales homogéneas en tres incógnitas $[\mathbf{f}(\mathbf{x}^1) - \mathbf{y}^1]_i$. Como el determinante de este sistema es distinto de cero,

$J\mathbf{f}(\mathbf{x}^1) \neq 0$, el sistema tiene la solución única $[\mathbf{f}(\mathbf{x}^1)-\mathbf{y}^1]_i = 0$ para cada i (problema 2, pág. 66). Es decir, $\mathbf{f}(\mathbf{x}^1) = \mathbf{y}^1$ para algún punto $\mathbf{x}^1 \in \mathscr{S}$.

Nota. Los resultados del lema 6.1 y el teorema 6.3 se verifican cuando reemplazamos R^3 por R^n. Las pruebas son esencialmente las mismas. Sin embargo, para una función de R^n en R^n el jacobiano es el determinante de una matriz $n \times n$, un concepto que nosotros hemos discutido solamente en los casos $n = 2$ y $n = 3$.

6.4 Lema. *Sea $\mathscr{R}$ un intervalo en R^3 con lados de longitudes $a \leqslant b \leqslant c$. Entonces existen vecindades $\mathscr{S}_1, \ldots, \mathscr{S}_{mn}$ de diámetro $\sqrt{3}a$ tales que*

$$\mathscr{S}_1 \cup \ldots \cup \mathscr{S}_{mn} \supset \mathscr{R} \quad y \quad \sum_{i=1}^{mn} V(\mathscr{S}_i) < 2\sqrt{3}\pi V(\mathscr{R}).$$

Prueba. Por la propiedad arquimediana de los números reales (vol. I, pág. 429, problema 3) existen enteros m y n tales que

$$(m-1)a < b \leqslant ma,$$
$$(n-1)a < c \leqslant na.$$

De donde $\mathscr{R}$ puede ser cubierto por mn cubos de volumen a^3 cada uno, y cada uno de estos mn cubos puede cubrirse por una vecindad de diámetro $\sqrt{3}a$ y volumen $\frac{\sqrt{3}}{2}\pi a^2$. Como $b \leqslant c$, tenemos $m \leqslant n$. Consideramos estos tres casos.

Caso 1. Si $2 \leqslant m \leqslant n$, entonces

$$mn\frac{\sqrt{3}}{2}\pi a^3 \leqslant 2(m-1)\,2(n-1)\frac{\sqrt{3}}{2}\pi a^3 < 2\sqrt{3}\pi abc.$$

Caso 2. Si $1 = m < 2 \leqslant n (b = a)$, entonces

$$n\frac{\sqrt{3}}{2}\pi a^3 \leqslant 2(n-1)\frac{\sqrt{3}}{2}\pi a^3 < \sqrt{3}\pi abc < 2\sqrt{3}\pi abc.$$

Caso 3. Si $m = n = 1$ $(a = b = c)$, entonces

$$\frac{\sqrt{3}}{2}\pi a^3 < 2\sqrt{3}\pi a^3 = 2\sqrt{3}\pi\, abc.$$

6.5 Teorema. *Sea $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$ un conjunto en R^3 de volumen cero y sea $\mathbf{f}$ una transformación de R^3 en R^3 de clase C^1 sobre un conjunto abierto $\mathscr{G}$ que contiéne $[\mathbf{a}, \mathbf{b}]$. Entonces la imagen $\mathbf{f}(\mathscr{E})$ de $\mathscr{E}$ bajo $\mathbf{f}$ tiene volumen cero.*

PRUEBA. Como $\mathbf{f}$ es de clase C^1 sobre $\mathscr{G}$, las derivadas parciales $D_j f_i$ son continuas sobre $\mathscr{G}$ y

$$\|D\mathbf{f}\| = \left[\sum_{i=1}^{3}\sum_{j=1}^{3}(D_j f_i)^2\right]^{1/2}$$

es continua sobre $\mathscr{G}$. Por tanto $\|D\mathbf{f}\|$ es acotada sobre $[\mathbf{a}, \mathbf{b}]$; es decir, para algún número K, $\|D\mathbf{f}(\mathbf{x})\| < K$ para todo $\mathbf{x}\in[\mathbf{a}, \mathbf{b}]$. Por el teorema 4.6, pág. 264, $\mathbf{f}$ es diferenciable sobre $[\mathbf{a}, \mathbf{b}]$ y para cada $\varepsilon > 0$ hay una $\delta > 0$ tal que $\mathbf{x}$. $\mathbf{y}\in[\mathbf{a}, \mathbf{b}]$ y $|\mathbf{y}-\mathbf{x}| < \delta$ implica

$$\begin{aligned}|\mathbf{f}(\mathbf{y})-\mathbf{f}(\mathbf{x})| &= |D\mathbf{f}(\mathbf{x})(\mathbf{y}-\mathbf{x})+\Phi(\mathbf{x};\mathbf{y}-\mathbf{x})(\mathbf{y}-\mathbf{x})| \\ &\leqslant [\|D\mathbf{f}(\mathbf{x})\| + \|\Phi(\mathbf{x};\mathbf{y}-\mathbf{x})\|]\,|\mathbf{y}-\mathbf{x}| < M|\mathbf{y}-\mathbf{x}|\end{aligned}$$

donde $M = K+\varepsilon$. Así pues, si $\mathscr{S}$ es una vecindad de radio $r < \delta$ y centro $\mathbf{x}$, entonces $\mathbf{f}$ transforma $\mathscr{S} \cap [\mathbf{a}, \mathbf{b}]$ en el interior de una vecindad de radio Mr. De aquí que $\bar{V}(\mathbf{f}(\mathscr{S} \cap [\mathbf{a}, \mathbf{b}])) \leqslant \frac{4}{3}\pi(Mr)^3 = M^3 V(\mathscr{S})$.

Pero como $\mathscr{E}$ tiene un volumen cero, para cada $\varepsilon > 0$ hay una partición P de $[\mathbf{a}, \mathbf{b}]$ tal que la unión $\mathscr{B}$ de aquellos subintervalos determinados por P que contienen puntos de $\bar{\mathscr{E}}$ tiene volumen $V(\mathscr{B}) < \varepsilon$. Sea P' un refinamiento de P de norma menor que δ. Según el lema 6.4, cada intervalo $\mathscr{R}$ determinado por la partición P' puede ser cubierto por una colección finita de vecindades esféricas de diámetro menor que $\sqrt{3}\,\delta$ la suma de cuyos volúmenes es menor que $2\sqrt{3}\pi V(\mathscr{R})$. De donde $\mathscr{B}$ puede cubrirse con una colección finita de vecindades $\mathscr{S}_1, \ldots, \mathscr{S}_N$ de diámetro menor que $\sqrt{3}\delta$, la suma de cuyos volúmenes es menor que

$$2\sqrt{3}\pi V(\mathscr{B}) < 2\sqrt{3}\pi\varepsilon .$$

Sea $\mathscr{A} = [\mathbf{a}, \mathbf{b}] \cap \bigcup_{i=1}^{N} \mathscr{S}_i$. Entonces $\mathscr{E} \subset \mathscr{B} \subset \mathscr{A}$ y $\mathbf{f}(\mathscr{E}) \subset \mathbf{f}(\mathscr{A})$. Por tanto

$$\begin{aligned}\bar{V}(\mathbf{f}(\mathscr{E})) \leqslant \bar{V}(\mathbf{f}(\mathscr{A})) &\leqslant \sum_{i=1}^{N} \bar{V}(\mathbf{f}(\mathscr{S}_i \cap [\mathbf{a}, \mathbf{z}])) \leqslant M^3 \sum_{i=1}^{N} V(\mathscr{S}_i) \\ &\leqslant M^3\, 2\sqrt{3}\pi V(\mathscr{B}) < M^3\, 2\sqrt{3}\pi\varepsilon .\end{aligned}$$

Pero $\varepsilon > 0$ es arbitrario, de modo que $\mathbf{f}(\mathscr{E})$ tiene volumen exterior cero y por consiguiente volumen cero.

6.6 Corolario. *Sea $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$ un conjunto en R^3 que tiene volumen, y sea $\mathbf{f}$ una transformación de R^3 en R^3 de clase C^1 sobre un conjunto abierto $\mathscr{G}$ que contiene a $[\mathbf{a}, \mathbf{b}]$. Si el jacobiano $J\mathbf{f}(\mathbf{x}) \neq 0$ para todo $\mathbf{x}\in\mathscr{G}$, entonces la imagen $\mathbf{f}(\mathscr{E})$ de $\mathscr{E}$ bajo $\mathbf{f}$ tiene volumen.*

PRUEBA. Primero demostramos que la frontera de la imagen de $\mathscr{E}$ está contenida en la imagen de la frontera de $\mathscr{E}$: $\mathbf{f}(\mathscr{E})_b \subset \mathbf{f}(\mathscr{E}_b)$. Como $\bar{\mathscr{E}}$ es un

conjunto cerrado y acotado y $\mathbf{f}$ es continua sobre $\bar{\mathscr{E}}$, $\mathbf{f}(\bar{\mathscr{E}})$ es cerrado según el teorema 7.8, pág. 478. Así pues, como $\mathbf{f}(\mathscr{E}) \subset \mathbf{f}(\bar{\mathscr{E}})$ y $\mathbf{f}(\bar{\mathscr{E}})$ es cerrado,

$$\mathbf{f}(\mathscr{E})_b \subset \mathbf{f}(\bar{\mathscr{E}}) = \mathbf{f}(\mathscr{E}_i \cup \mathscr{E}_b) = \mathbf{f}(\mathscr{E}_i) \cup \mathbf{f}(\mathscr{E}_b)$$

Según el teorema 6.3, $\mathbf{f}(\mathscr{E}_i)$ es abierto, de modo que $\mathbf{f}(\mathscr{E}_i) = \mathbf{f}(\mathscr{E}_i)_i \subset \mathbf{f}(\mathscr{E})_i$. Luego $\mathbf{f}(\mathscr{E}_i) \cap \mathbf{f}(\mathscr{E})_b \subset \mathbf{f}(\mathscr{E})_i \cap \mathbf{f}(\mathscr{E})_b = \varnothing$ y como $\mathbf{f}(\mathscr{E})_b \subset \mathbf{f}(\mathscr{E}_i) \cup \mathbf{f}(\mathscr{E}_b)$ tenemos $\mathbf{f}(\mathscr{E})_b \subset \mathbf{f}(\mathscr{E}_b)$.

Como $\mathscr{E}$ tiene volumen, $\mathscr{E}_b$ tiene volumen cero (teorema 4.17, pág. 337, generalizado a R^3). Según el teorema 6.5, $\mathbf{f}(\mathscr{E}_b)$ tiene volumen cero y, por tanto, $\mathbf{f}(\mathscr{E})_b$ tiene volumen cero. De donde $\mathbf{f}(\mathscr{E})$ tiene volumen, según el corolario 14.14, pág. 379.

En el próximo lema demostraremos que si $\mathbf{f}$ es una transformación de R^3 en R^3 de clase C^1 sobre un conjunto abierto $\mathscr{G}$ y si $J\mathbf{f}(\mathbf{x}^c) \neq 0$ para un punto $\mathbf{x}^c \in \mathscr{G}$, entonces la imagen $\mathbf{f}(\mathscr{R})$ de cualquier intervalo suficientemente pequeño $\mathscr{R}$ que no difiera demasiado de un cubo y tenga centro en $\mathbf{x}^c$ tiene volumen $V(\mathbf{f}(\mathscr{R}))$ aproximadamente igual a $V(\mathscr{R})\, J\mathbf{f}(\mathbf{x}^c)$.

6.7 Lema. *Sea* $\mathbf{f}$ *una transformación de clase* C^1 *de* R^3 *en* R^3 *definida sobre un conjunto abierto* $\mathscr{G}$ *y sea el jacobiano* $J\mathbf{f}(\mathbf{x}^c)$ *distinto de cero en un punto* $\mathbf{x}^c \in \mathscr{G}$. *Para cada* $\varepsilon > 0$ *hay una* $\delta > 0$ *tal que si* $\mathscr{R}$ *es un intervalo en* $\mathscr{G}$ *con centro en* $\mathbf{x}^c$, *diámetro* $2r < \delta$, *y lados de longitud* $2a_k$ *tales que* $0 < a \leqslant a_k \leqslant 2a$ *para todo* $k = 1, 2, 3$, *donde* $a = \text{mín}\,\{a_1, a_2, a_3\}$, *entonces*

$$|V(\mathbf{f}(\mathscr{R})) - V(\mathscr{R})|\, J\mathbf{f}(\mathbf{x}^c)|\,| < \varepsilon V(\mathscr{R})\, M(\mathbf{x}^c)$$

donde $M(\mathbf{x}^c)$ *es independiente de* $\mathscr{R}$.

PRUEBA. Como $\mathbf{f}$ es diferenciable en $\mathbf{x}^c$,

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}^c) + D\mathbf{f}(\mathbf{x}^c)(\mathbf{x}-\mathbf{x}^c) + \Phi(\mathbf{x}^c; \mathbf{x}-\mathbf{x}^c)(\mathbf{x}-\mathbf{x}^c) \text{ para } \mathbf{x} \in \mathscr{G}$$

donde $\lim_{\mathbf{x}\to\mathbf{x}^c} \Phi(\mathbf{x}^c; \mathbf{x}-\mathbf{x}^c) = O$. Por tanto, para cada $\varepsilon > 0$ hay una $\delta > 0$ tal que $\|\Phi(\mathbf{x}^c; \mathbf{x}-\mathbf{x}^c)\| < \varepsilon$ siempre que $\mathbf{x} \in \mathscr{S}(\mathbf{x}^c; \delta)$. Como $\mathscr{G}$ es abierto, podemos tomar una $\delta > 0$ tan pequeña que $\mathscr{S}(\mathbf{x}^c; \delta) \subset \mathscr{G}$. Sea $\mathbf{g}^c$ la transformación definida sobre $\mathscr{G}$ por la regla de correspondencia

$$\mathbf{g}^c(\mathbf{x}) = \mathbf{f}(\mathbf{x}^c) + D\mathbf{f}(\mathbf{x}^c)(\mathbf{x}-\mathbf{x}^c).$$

Si $\mathscr{R}$ es el intervalo en el enunciado del lema con centro $\mathbf{x}^c$ y diámetro $2r < \delta$, entonces para cada $\mathbf{x} \in \mathscr{R}$

$$|\mathbf{f}(\mathbf{x}) - \mathbf{g}^c(\mathbf{x})| = |\Phi(\mathbf{x}^c; \mathbf{x}-\mathbf{x}^c)(\mathbf{x}-\mathbf{x}^c)| \leqslant \|\Phi(\mathbf{x}^c; \mathbf{x}-\mathbf{x}^c)\|\, |\mathbf{x}-\mathbf{x}^c| < \varepsilon r.$$

El intervalo $\mathscr{R}$ es el paralelepípedo rectangular

$$\mathscr{R} = \{\mathbf{x}^c + \sum_{k=1}^{3} t_k a_k \mathbf{u}_k \mid t_k \in [-1, 1]\}$$

y la imagen de $\mathscr{R}$ bajo $\mathbf{g}^c$ es el paralelepípedo

$$\mathbf{g}^c(\mathscr{R}) = \{\mathbf{f}(\mathbf{x}^c) + \sum_{k=1}^{3} t_k a_k \mathbf{b}_k \mid t_k \in [-1, 1]\}$$

donde $\mathbf{b}_k = D_k\mathbf{f}(\mathbf{x}^c)$.

Demostraremos que $\mathbf{f}(\mathscr{R})$ se encuentra en el interior de un paralelepípedo $\mathscr{P}_1$ que es semejante a $\mathbf{g}^c(\mathscr{R})$ y tiene caras a una distancia εr fuera de las de $\mathbf{g}^c(\mathscr{R})$ y que $\mathbf{f}(\mathscr{R})$ cubre un paralelepípedo $\mathscr{P}_2$ que es análogo a $\mathbf{g}^c(\mathscr{R})$ y tiene caras a una distancia εr hacia el interior de las de $\mathbf{g}^c(\mathscr{R})$:

$$\mathscr{P}_1 = \{\mathbf{f}(\mathbf{x}^c) + \sum_{k=1}^{3} t_k \mathbf{b}_k(a_k + \varepsilon r d_k) \mid t_k \in [-1, 1]\}$$

$$\mathscr{P}_2 = \{\mathbf{f}(\mathbf{x}^c) + \sum_{k=1}^{3} t_k \mathbf{b}_k(a_k - \varepsilon r d_k) \mid t_k \in [-1, 1]\}$$

donde $d_1 = \dfrac{|\mathbf{b}_2 \times \mathbf{b}_3|}{|J\mathbf{f}(\mathbf{x}^c)|}$, $d_2 = \dfrac{|\mathbf{b}_3 \times \mathbf{b}_1|}{|J\mathbf{f}(\mathbf{x}^c)|}$, y $d_3 = \dfrac{|\mathbf{b}_1 \times \mathbf{b}_2|}{|J\mathbf{f}(\mathbf{x}^c)|}$. Como $\dfrac{r}{a_k} \leqslant \dfrac{r}{a} < \dfrac{1}{a}\sqrt{a^2+(2a)^2+(2a)^2} = 3$, y d_k es independiente de ε, sin pérdida de generalidad podemos suponer que $\varepsilon > 0$ es tan pequeña que $1 > 3\varepsilon d_k \geqslant \dfrac{\varepsilon r d_k}{a_k}$ para $k = 1, 2, 3$. Pero $\mathbf{x}\in\mathscr{R}$ implica $|\mathbf{f}(\mathbf{x}) - \mathbf{g}^c(\mathbf{x})| < \varepsilon r$ y por tanto $\mathbf{f}(\mathbf{x})\in\mathscr{P}_1$. De donde $\mathbf{f}(\mathbf{x})\subset\mathscr{P}_1$. Por otra parte $\mathbf{g}^c$ transforma la frontera de $\mathscr{R}$ sobre la frontera de $\mathbf{g}^c(\mathscr{R})$ de modo que si $\mathbf{x}\in\mathscr{R}_b$, $\mathbf{f}(\mathbf{x})$ está a una distancia menor que εr de la frontera de $\mathbf{g}^c(\mathscr{R})$ y, por tanto, se encuentra fuera de $\mathscr{P}_2$. Luego $\mathbf{f}(\mathscr{R}_b)$ se encuentra fuera de $\mathscr{P}_2$. Pero $\mathbf{f}(\mathbf{x}^c)\in\mathscr{P}_2$ y si cualquier punto de $\mathscr{P}_2$ no fuera un punto de $\mathbf{f}(\mathscr{R})$, habría puntos de $\mathbf{f}(\mathscr{R})_b$ en $\mathscr{P}_2$ lo que es imposible, pues, como demostramos en la prueba del corolario 6.6 $\mathbf{f}(\mathscr{R})_b \subset \mathbf{f}(\mathscr{R}_b)$, y éste está fuera de $\mathscr{P}_2$. Por tanto $\mathscr{P}_2 \subset \mathbf{f}(\mathscr{R}) \subset \mathscr{P}_1$. De donde $V(\mathscr{P}_2) \leqslant V(\mathbf{f}(\mathscr{R})) \leqslant V(\mathscr{P}_1)$. Como $\dfrac{r}{a_k} < 3$ y $3\varepsilon d_k < 1$ para $k = 1, 2, 3$, tenemos

$$V(\mathscr{P}_1) = 8a_1a_2a_3|[\mathbf{b}_1\mathbf{b}_2\mathbf{b}_3]| \prod_{k=1}^{3}\left(1 + \frac{\varepsilon r d_k}{a_k}\right) < V(\mathscr{R})|J\mathbf{f}(\mathbf{x}^c)| \prod_{k=1}^{3}(1 + 3\varepsilon d_k)$$

y

$$V(\mathscr{P}_2) = 8a_1a_2a_3|[\mathbf{b}_1\mathbf{b}_2\mathbf{b}_3]| \prod_{k=1}^{3}\left(1 - \frac{\varepsilon r d_k}{a_k}\right) > V(\mathscr{R})|J\mathbf{f}(\mathbf{x}^c)| \prod_{k=1}^{3}(1 - 3\varepsilon d_k)$$

Por tanto

$$-\varepsilon V(\mathscr{R})M(\mathbf{x}^c) < V(\mathscr{P}_2) - V(\mathscr{R})\,|J\mathbf{f}(\mathbf{x}^c)| \leqslant V(\mathbf{f}(\mathscr{R})) - V(\mathscr{R})\,|J\mathbf{f}(\mathbf{x}^c)|$$
$$\leqslant V(\mathscr{P}_1) - V(\mathscr{R})\,|J\mathbf{f}(\mathbf{x}^c)| < \varepsilon V(\mathscr{R})M(\mathbf{x}^c)$$

y

$$|V(\mathbf{f}(\mathscr{R}))-V(\mathscr{R})\,|J\mathbf{f}(\mathbf{x}^c)|\,| < \varepsilon V(\mathscr{R})\,M(\mathbf{x}^c)$$

donde

$$M(\mathbf{x}^c) = [3(d_1+d_2+d_3)+9\,\varepsilon(d_1 d_2+d_2 d_3+d_3 d_1)+27\,\varepsilon^2 d_1 d_2 d_3]\,|J\mathbf{f}(\mathbf{x}^c)|.$$

En el próximo lema probaremos que $V \circ \mathbf{f}$ es uniformemente diferenciable sobre todo intervalo $[\mathbf{a}, \mathbf{b}] \subset \mathscr{G}$.

6.8 Lema. *Sea una transformación* $\mathbf{f}$ *de* R^3 *en* R^3 *de clase* C^1 *sobre un conjunto abierto* $\mathscr{G}$ *y sea el jacobiano* $J\mathbf{f}(\mathbf{x})$ *distinto de cero para todo* $\mathbf{x}\in\mathscr{G}$. *Si* $[\mathbf{a}, \mathbf{b}] \subset \mathscr{G}$, *entonces para* $\mathbf{x}\in[\mathbf{a}, \mathbf{b}]$

$$\lim_{\mathscr{E}\to\mathbf{x}} \frac{V(\mathbf{f}(\mathscr{E}))}{V(\mathscr{E})} = |J\mathbf{f}(\mathbf{x})|$$

y el límite es uniforme sobre $[\mathbf{a}, \mathbf{b}]$.

PRUEBA. Sea $2\delta_1 = d([\mathbf{a}, \mathbf{b}], \mathscr{G}_b)$. Entonces $\delta_1 > 0$ según el teorema 4.1, pág. 411. Sea $\mathscr{F} = \{\mathbf{x} \mid d(\mathbf{x}, [\mathbf{a}, \mathbf{b}]) \leqslant \sqrt{3}\,\delta_1\}$. Entonces $\mathscr{F}$ es un subconjunto cerrado de $\mathscr{G}$. Como $\mathbf{f}$ es de clase C^1 sobre $\mathscr{G}$, $\mathbf{f}$ es diferenciable en todo $\mathbf{x}^0\in\mathscr{F}$ y

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}^0)+D\mathbf{f}(\mathbf{x}^0)\,(\mathbf{x}-\mathbf{x}^0)+\Phi(\mathbf{x}^0;\mathbf{x}-\mathbf{x}^0)\,(\mathbf{x}-\mathbf{x}^0)$$

donde $\lim_{\mathbf{x}\to\mathbf{x}^0} \Phi(\mathbf{x}^0;\mathbf{x}-\mathbf{x}^0) = 0$ uniformemente sobre $\mathscr{F}$ (teorema 4.6, pág. 263). Luego, para cualquier $\varepsilon > 0$ que tomemos hay una $\delta_2 > 0$ tal que

$$\|\Phi(\mathbf{x}^0;\mathbf{x}-\mathbf{x}^0)\| < \varepsilon$$

siempre que $\mathbf{x}^0, \mathbf{x}\in\mathscr{F}$ y $|\mathbf{x}-\mathbf{x}^0| < \delta_2$. Como $\mathbf{f}$ es de clase C^1 sobre $\mathscr{G}$, según el teorema 7.6, pág. 478, el jacobiano $J\mathbf{f}$ es uniformemente continuo sobre $\mathscr{F}$. Hay pues una $\delta_3 > 0$ tal que $|J\mathbf{f}(\mathbf{x})-J\mathbf{f}(\mathbf{x}^0)| < \varepsilon$ siempre que $\mathbf{x}^0, \mathbf{x}\in\mathscr{F}$ y $|\mathbf{x}-\mathbf{x}^0| < \delta_3$. Además, de acuerdo con el teorema 7.7, pág. 478, la continuidad de $J\mathbf{f}$ sobre $\mathscr{F}$ implica que $J\mathbf{f}$ está acotado en $\mathscr{F}$ y existe un número J tal que $|J\mathbf{f}(\mathbf{x})| \leqslant J$ para todo $\mathbf{x}\in\mathscr{F}$.

Tómese $\mathbf{x}^0\in[\mathbf{a}, \mathbf{b}]$ y sea $\mathscr{I}$ el intervalo (cubo) con centro $\mathbf{x}^0$ y lados de longitud 2δ donde $\delta = \text{mín}\{\delta_1, \delta_2, \delta_3\}$. Entonces $\mathscr{S}(\mathbf{x}^0;\delta)\subset\mathscr{I}\subset\mathscr{F}\subset\mathscr{G}$. Demostraremos que si $\mathscr{E}$ es un subconjunto de $\mathscr{S}(\mathbf{x}^0;\delta)$ que contiene a $\mathbf{x}^0$ y tiene volumen positivo, entonces

6.9
$$\left|\frac{V(\mathbf{f}(\mathscr{E}))}{V(\mathscr{E})} - |J\mathbf{f}(\mathbf{x}^0)|\right| < \varepsilon N$$

donde N es independiente de $\mathscr{E}$. Ahora bien, $\mathscr{E} \subset \mathscr{S}(\mathbf{x}^0;\delta) \subset \mathscr{I}$ y como $\mathscr{E}$

tiene volumen, hay una partición P de $\mathscr{I}$ tal que $V(\mathscr{B}) - V(\mathscr{A}) < \varepsilon V(\mathscr{E})$ donde $\mathscr{A}$ es la unión de aquellos subintervalos determinados por P que están contenidos en $\mathscr{E}_i$ y $\mathscr{B}$ es la unión de aquellos subintervalos que contienen puntos de $\bar{\mathscr{E}}$. Entonces

6.10 $$V(\mathscr{A}) \leqslant V(\mathscr{E}) \leqslant V(\mathscr{B})$$

y

6.11 $$V(\mathbf{f}(\mathscr{A})) \leqslant V(\mathbf{f}(\mathscr{E})) \leqslant V(\mathbf{f}(\mathscr{B})).$$

Por refinamiento de la partición P, si es necesario, podemos obtener una partición P' tal que cada subintervalo $\mathscr{R}_j$ tenga lados de longitud $2a_k{}^j$ tales que $0 < a^j \leqslant a_k{}^j < 2a^j$ donde $a^j = \text{mín}\,\{a_1{}^j, a_2{}^j, a_3{}^j\}$. Entonces, por el lema 6.7

$$\begin{aligned}|V(\mathbf{f}(\mathscr{R}_j)) - V(\mathscr{R}_j)\,|J\mathbf{f}(\mathbf{x}^0)||\ &\leqslant |V(\mathbf{f}(\mathscr{R}_j)) - V(\mathscr{R}_j)\,|J\mathbf{f}(\mathbf{x}^j)||\\ &\quad + V(\mathscr{R}_j)\,|J\mathbf{f}(\mathbf{x}^j) - J\mathbf{f}(\mathbf{x}^0)| < \varepsilon V(\mathscr{R}_j)\,[M(\mathbf{x}^j)+1],\end{aligned}$$

donde $\mathbf{x}^j$ es el centro de $\mathscr{R}_j$. Como $\mathbf{f}$ es de clase C^1 sobre $\mathscr{G}$, $M(\mathbf{x})$ es continua sobre $\mathscr{I}$ y por el teorema 7.7, pág. 478, existe un número M tal que $M(\mathbf{x}) + 1 \leqslant M$ para todo $\mathbf{x} \in \mathscr{I}$. De donde

$$\begin{aligned}|V(\mathbf{f}(\mathscr{B})) - V(\mathscr{B})\,|J\mathbf{f}(\mathbf{x}^0)|| &\leqslant \sum_{\mathscr{R}_j \subset \mathscr{B}} |V(\mathbf{f}(\mathscr{R}_j)) - V(\mathscr{R}_j)\,|J\mathbf{f}(\mathbf{x}^0)||\\ &< \varepsilon \sum_{\mathscr{R}_j \subset \mathscr{B}} V(\mathscr{R}_j) M = \varepsilon V(\mathscr{B}) M\end{aligned}$$

o bien,

$$\left| \frac{V(\mathbf{f}(\mathscr{B}))}{V(\mathscr{B})} - |J\mathbf{f}(\mathbf{x}^0)| \right| < \varepsilon M.$$

Análogamente podemos obtener

$$\left| \frac{V(\mathbf{f}(\mathscr{A}))}{V(\mathscr{A})} - |J\mathbf{f}(\mathbf{x}^0)| \right| < \varepsilon M.$$

De las desigualdades 6.10 y 6.11, se sigue que

$$\frac{V(\mathbf{f}(\mathscr{A}))}{V(\mathscr{B})} \leqslant \frac{V(\mathbf{f}(\mathscr{E}))}{V(\mathscr{E})} \leqslant \frac{V(\mathbf{f}(\mathscr{B}))}{V(\mathscr{A})}$$

y

$$\frac{V(\mathbf{f}(\mathscr{A}))}{V(\mathscr{B})} \leqslant \frac{V(\mathbf{f}(\mathscr{B}))}{V(\mathscr{B})} \leqslant \frac{V(\mathbf{f}(\mathscr{B}))}{V(\mathscr{A})}$$

de modo que si probamos que los miembros extremos de estas desigualdades difieren en menos de algún múltiplo de ε independientemente de $\mathscr{E}$, entonces

lo mismo ocurrirá con los dos términos medios y seremos capaces de establecer 6.9. Pero como

$$V(\mathscr{B})-V(\mathscr{A}) < \varepsilon V(\mathscr{E}) \leqslant \varepsilon V(\mathscr{B}),$$

tenemos

$$(1-\varepsilon)\, V(\mathscr{B}) < V(\mathscr{A}).$$

Suponiendo $\varepsilon < 1$, tenemos

$$\left|\frac{V(\mathbf{f}(\mathscr{E}))}{V(\mathscr{E})}-\frac{V(\mathbf{f}(\mathscr{B}))}{V(\mathscr{B})}\right| \leqslant \frac{V(\mathbf{f}(\mathscr{B}))}{V(\mathscr{A})}-\frac{V(\mathbf{f}(\mathscr{A}))}{V(\mathscr{B})}$$

$$=\frac{V(\mathscr{B})}{V(\mathscr{A})}\,\frac{V(\mathbf{f}(\mathscr{B}))}{V(\mathscr{B})}-\frac{V(\mathscr{A})}{V(\mathscr{B})}\,\frac{V(\mathbf{f}(\mathscr{A}))}{V(\mathscr{A})}$$

$$<\frac{1}{1-\varepsilon}\left[|J\mathbf{f}(\mathbf{x}^0)|+\varepsilon M\right]-(1-\varepsilon)\left[|J\mathbf{f}(\mathbf{x}^0)|-\varepsilon M\right]$$

$$=\frac{2\varepsilon-\varepsilon^2}{1-\varepsilon}\,|J\mathbf{f}(\mathbf{x}^0)|+\frac{2-2\varepsilon+\varepsilon^2}{1-\varepsilon}\,\varepsilon M$$

$$\leqslant \varepsilon\left[\frac{2-\varepsilon}{1-\varepsilon}J+\frac{2-2\varepsilon+\varepsilon^2}{1-\varepsilon}M\right]$$

donde J es una cota superior para $|J\mathbf{f}(\mathbf{x})|$ sobre $[\mathbf{a}, \mathbf{b}]$. De donde

$$\left|\frac{V(\mathbf{f}(\mathscr{E}))}{V(\mathscr{E})}-|J\mathbf{f}(\mathbf{x}^0)|\right| \leqslant \left|\frac{V(\mathbf{f}(\mathscr{E}))}{V(\mathscr{E})}-\frac{V(\mathbf{f}(\mathscr{B}))}{V(\mathscr{B})}\right|+\left|\frac{V(\mathbf{f}(\mathscr{B}))}{V(\mathscr{B})}-|J\mathbf{f}(\mathbf{x}^0)|\right| < \varepsilon N$$

donde $N=\dfrac{1}{1-\varepsilon}\left[(2-\varepsilon)J+(3-3\varepsilon+\varepsilon^2)M\right]$ es independiente de $\mathscr{E}$ y $\mathbf{x}^0\in[\mathbf{a},\mathbf{b}]$. Esto completa la prueba del lema.

De acuerdo con el lema 6.8,

6.12 $$V(\mathbf{f}(\mathscr{E})) = |J\mathbf{f}(\mathbf{x})|\; V(\mathscr{E})+\Psi(\mathbf{x};\mathscr{E})\; V(\mathscr{E})$$

donde $\lim\limits_{\mathscr{E}\to\mathbf{x}} \Psi(\mathbf{x};\mathscr{E}) = 0$ uniformemente para $\mathbf{x}\in[\mathbf{a}, \mathbf{b}]$. Sea $\mathscr{G}$, una función de conjunto definida sobre una familia $\mathfrak{N}$ de conjuntos en R^3 que tienen volumen, diferenciable sobre un conjunto $\mathscr{S}\subset R^3$. Entonces

6.13 $$G(\mathscr{F}) = DG(\mathbf{y})\; V(\mathscr{F})+\Phi(\mathbf{y};\mathscr{F})\; V(\mathscr{F})$$

donde $DG(\mathbf{y})$ es independiente de $\mathscr{F}$ y $\lim\limits_{\mathscr{F}\to\mathbf{y}} \Phi(\mathbf{y};\mathscr{F}) = 0$ para todo $\mathbf{y}\in\mathscr{S}$. Sea $\mathbf{f}$, una transformación de R^3 en R^3, de clase C^1 sobre un conjunto

abierto $\mathscr{G}$ y sea $J\mathbf{f}(\mathbf{x})$, el jacobiano de $\mathbf{f}$, distinto de cero para todo $\mathbf{x}\in\mathscr{G}$. Definamos la función de conjunto $F = G \circ \mathbf{f}$ por la regla de correspondencia

$$F(\mathscr{E}) = G(\mathbf{f}(\mathscr{E}))$$

donde $\mathscr{E}$ es un subconjunto de $\mathscr{G}$ que tiene volumen y $\mathbf{f}(\mathscr{E})\in\mathfrak{D}_G$. Supongamos que $\mathbf{x}\in\mathscr{G}$ y $\mathbf{f}(\mathbf{x})\in\mathscr{S}$. Entonces, por las ecuaciones 6.12 y 6.13, tenemos

6.14
$$\begin{aligned} F(\mathscr{E}) &= G(\mathbf{f}(\mathscr{E})) = [DG(\mathbf{f}(\mathbf{x}))+\Phi(\mathbf{f}(\mathbf{x});\mathbf{f}(\mathscr{E}))]\, V(\mathbf{f}(\mathscr{E})) \\ &= [DG(\mathbf{f}(\mathbf{x}))+\Phi(\mathbf{f}(\mathbf{x});\mathbf{f}(\mathscr{E}))]\,[|J\mathbf{f}(\mathbf{x})|+\Psi(\mathbf{x};\mathscr{E})]\, V(\mathscr{E}) \\ &= DG(\mathbf{f}(\mathbf{x}))\,|J\mathbf{f}(\mathbf{x})|\, V(\mathscr{E})+\Theta(\mathbf{x};\mathscr{E})\, V(\mathscr{E}) \end{aligned}$$

donde $DG(\mathbf{f}(\mathbf{x}))\,|J\mathbf{f}(\mathbf{x})|$ es independiente de $\mathscr{E}$ y

$$\lim_{\mathscr{E}\to\mathbf{x}} \Theta(\mathbf{x};\mathscr{E}) = \lim_{\mathscr{E}\to\mathbf{x}} \{\Phi(\mathbf{f}(\mathbf{x});\mathbf{f}(\mathscr{E}))\,[|J\mathbf{f}(\mathbf{x})|+\Psi(\mathbf{x};\mathscr{E})]+DG(\mathbf{f}(\mathbf{x}))\,\Psi(\mathbf{x};\mathscr{E})\}=0$$

Así pues, F es diferenciable en $\mathbf{x}$ y tenemos la regla de la cadena

6.15
$$DF(\mathbf{x}) = DG(\mathbf{f}(\mathbf{x}))\,|J\mathbf{f}(\mathbf{x})|.$$

Estamos ahora en posición de probar nuestro teorema sobre cambio de variable en las integrales triples.

6.16 Teorema. *Sea* $\mathbf{f}$ *una transformación de* R^3 *en* R^3 *de clase* C^1 *y univalente sobre un conjunto abierto* $\mathscr{G}$ *de jacobiano* $J\mathbf{f}(\mathbf{x}) \neq 0$ *para todo* $\mathbf{x}\in\mathscr{G}$ *y sea* g *una función de* R^3 *en* R *continua sobre* $\mathbf{f}(\mathscr{G})$. *Si* $[\mathbf{a}, \mathbf{b}] \subset \mathscr{G}$, $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$ *tiene volumen, y* $\mathscr{F} = \mathbf{f}(\mathscr{E})$, *entonces*

$$\int_{\mathscr{F}} g = \int_{\mathscr{F}} g(\mathbf{y})\,d\mathbf{y} = \int_{\mathscr{E}} (g\circ\mathbf{f})\,|J\mathbf{f}| = \int_{\mathscr{E}} g(\mathbf{f}(\mathbf{x}))\,|J\mathbf{f}(\mathbf{x})|\,d\mathbf{x}.$$

Prueba. Si g es una función constante, el teorema se sigue del lema 6.8 y del segundo teorema fundamental. Según el lema 6.8, $V \circ \mathbf{f}$ es uniformemente diferenciable sobre $[\mathbf{a}, \mathbf{b}]$ con

$$D[V(\mathbf{f}(\mathbf{x}))] = |J\mathbf{f}(\mathbf{x})| \quad \text{sobre } [\mathbf{a}, \mathbf{b}].$$

Como $\mathbf{f}$ es univalente, $V \circ \mathbf{f}$ es aditiva y monótona sobre el anillo de los subconjuntos de $[\mathbf{a}, \mathbf{b}]$ que tienen volumen. Si $\mathscr{F} = \mathbf{f}(\mathscr{E})$ y $\mathscr{E} \subset [\mathbf{a}, \mathbf{b}]$ tiene volumen, entonces $\mathscr{F}$ tiene volumen según el corolario 6.6 y el segundo teorema fundamental

$$V(\mathscr{F}) = V(\mathbf{f}(\mathscr{E})) = \int_{\mathscr{E}} |J\mathbf{f}(\mathbf{x})|\,d\mathbf{x}.$$

Definamos la función de conjunto $\mathscr{G}$, sobre el anillo $\mathfrak{R}$ de subconjuntos de $\mathbf{f}([\mathbf{a}, \mathbf{b}])$ que tienen volumen, por la regla de correspondencia

$$G(\mathscr{F}) = \int_{\mathscr{F}} g.$$

Entonces, para g una función constante,

$$G(\mathscr{F}) = \int_{\mathscr{F}} g = g \int_{\mathscr{F}} 1 = gV(\mathscr{F}) = g \int_{\mathscr{E}} |J\mathbf{f}(\mathbf{x})|\, d\mathbf{x} = \int_{\mathscr{E}} g\, |J\mathbf{f}(\mathbf{x})|\, d\mathbf{x}.$$

Si g es no constante, como g es continua sobre el conjunto cerrado y acotado $\mathbf{f}([\mathbf{a}, \mathbf{b}])$, g es acotada en él. Podemos suponer que g toma sólo valores positivos sobre $\mathbf{f}([\mathbf{a}, \mathbf{b}])$, puesto que en otro caso podríamos considerar a $(g-m+1)-(-m+1)$ donde $m = \text{mín}\,\{g(\mathbf{y}) \mid \mathbf{y} \in \mathbf{f}([\mathbf{a}, \mathbf{b}])\}$. Definamos la función de conjunto F por la regla de correspondencia

$$F(\mathscr{E}) = G(\mathbf{f}(\mathscr{E}))$$

donde $\mathscr{E}$ es un subconjunto de $[\mathbf{a}, \mathbf{b}]$ que tiene volumen y $\mathbf{f}(\mathscr{E}) \in \mathfrak{D}_G$. Por el primer teorema fundamental del cálculo, $DG(\mathbf{y}) = g(\mathbf{y})$ para todo $\mathbf{y} \in \mathbf{f}([\mathbf{a}, \mathbf{b}])$ y por la regla de la cadena (ecuación 6.15), F es diferenciable sobre $[\mathbf{a}, \mathbf{b}]$ con

$$DF(\mathbf{x}) = DG(\mathbf{f}(\mathbf{x}))\, |J\mathbf{f}(\mathbf{x})| = g(\mathbf{f}(\mathbf{x}))\, |J\mathbf{f}(\mathbf{x})|.$$

Como $\mathbf{f}$ es univalente, $V(\mathscr{E}_1 \cap \mathscr{E}_2) = 0$ implica que $V(\mathbf{f}(\mathscr{E}_1) \cap \mathbf{f}(\mathscr{E}_2)) = V(\mathbf{f}(\mathscr{E}_1 \cap \mathscr{E}_2)) = 0$ y por tanto F es aditiva ya que lo es G. Como se supone que g sólo toma valores positivos, F es monótona. De acuerdo con el primer teorema fundamental, G es uniformemente diferenciable sobre $\mathbf{f}([\mathbf{a}, \mathbf{b}])$ y, por tanto, la ecuación 6.14 implica que F es uniformemente diferenciable sobre $[\mathbf{a}, \mathbf{b}]$. Luego F satisface las condiciones del segundo teorema fundamental del cálculo e

$$\int_{\mathscr{F}} g = G(\mathscr{F}) = G(\mathbf{f}(\mathscr{E})) = F(\mathscr{E}) = \int_{\mathscr{E}} DF(\mathbf{x})\, d\mathbf{x} = \int_{\mathscr{E}} g(\mathbf{f}(\mathbf{x}))\, |J\mathbf{f}(\mathbf{x})|\, d\mathbf{x}.$$

Nota. Si $J\mathbf{f}(\mathbf{x}) = 0$ sobre un conjunto $\mathscr{F}$ de volumen cero y $\mathscr{E} - \mathscr{F}$ es la unión de un número finito de conjuntos que tienen volumen, entonces la fórmula del teorema 6.16 aún se verifica. Supongamos

$$\mathscr{E} = (\mathscr{E} \cap \mathscr{F}) \cup \mathscr{E}_1 \cup \mathscr{E}_2 \cup \ldots \cup \mathscr{E}_n$$

donde $\mathscr{E}_1, \mathscr{E}_2, \ldots, \mathscr{E}_n$ no se traslapan y tienen volumen y $\mathscr{E} \cap \mathscr{F}$ tiene volumen cero. Según el teorema 6.5, $V(\mathbf{f}(\mathscr{E} \cap \mathscr{F})) = 0$, y el lema 6.11, pág. 314, generalizado a R^3

$$\int_{\mathbf{f}(\mathscr{E} \cap \mathscr{F})} g = 0 = \int_{\mathscr{E} \cap \mathscr{F}} (g \circ \mathbf{f})\, |J\mathbf{f}|.$$

Por tanto

$$\int_{\mathbf{f}(\mathscr{E})} g = \int_{\mathbf{f}(\mathscr{E} \cap \mathscr{F})} g + \sum_{k=1}^{n} \int_{\mathbf{f}(\mathscr{E}_k)} g = \int_{\mathscr{E} \cap \mathscr{F}} (g \circ \mathbf{f})\, |J\mathbf{f}| + \sum_{k=1}^{n} \int_{\mathscr{E}_k} (g \circ \mathbf{f})\, |J\mathbf{f}|$$

$$= \int_{\mathscr{E}} (g \circ \mathbf{f})\, |J\mathbf{f}|.$$

6.17 Ejemplo. Sea $\mathbf{f}$ la transformación definida por la regla de correspondencia

$$(x, y, z) = \mathbf{f}(u, v, w) = (u^2+v, u-v, w)$$

y sea $\mathscr{E}$ el tetraedro con vértices $(0, 0, 0)$, $(0, 0, 1)$, $(0, 1, 0)$ y $(1, 0, 0)$. Encuéntrese el volumen de $\mathbf{f}(\mathscr{E})$.

SOLUCIÓN. La región $\mathscr{E}$ se muestra en la figura $2a$ y $\mathbf{f}(\mathscr{E})$ se muestra en la figura $2b$. La transformación $\mathbf{f}$ no es univalente en R^3, pero para $u \geqslant -\frac{1}{2}$, $\mathbf{f}$ es univalente y tiene inversa $\mathbf{f}^*$:

$$\mathbf{f}: \begin{cases} x = u^2+v \\ y = u-v \\ z = w \end{cases} \qquad \mathbf{f}^*: \begin{cases} u = -\frac{1}{2}+\sqrt{x+y+\frac{1}{4}} \\ v = -y-\frac{1}{2}+\sqrt{x+y+\frac{1}{4}} \\ w = z. \end{cases}$$

El conjunto $\mathscr{E}$ se encuentra en la región donde $u \geqslant 0$ y está limitada por los planos de coordenadas y el plano $u+v+w = 1$. El plano $w = 0$ s

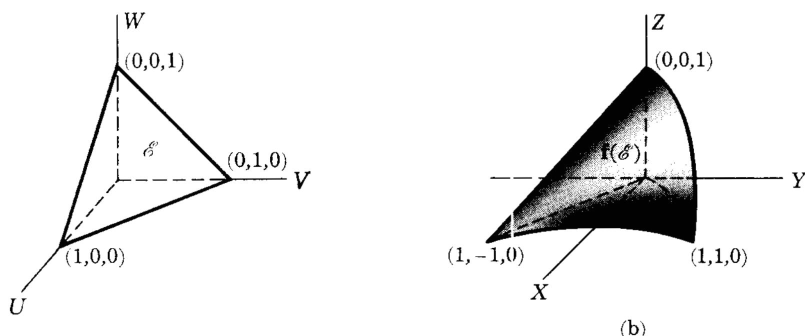


FIGURA 2

transforma en el plano $z = 0$; el plano $v = 0$ se transforma en el cilindro parabólico $x = y^2$; el plano $u = 0$ se transforma en el plano $x+y = 0$; y el plano $u+v+w = 1$ se transforma en la superficie $z-y+2\sqrt{x+y+\frac{1}{4}} = 2$. Tenemos

$$J\mathbf{f}(u, v, w) = \begin{vmatrix} 2u & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -2u-1$$

y

$$V(\mathbf{f}(\mathscr{E})) = \int_{\mathbf{f}(\mathscr{E})} 1 = \int_{\mathscr{E}} |J\mathbf{f}| = \int_0^1 \int_0^{1-u} \int_0^{1-u-v} (2u+1)\, dw\, dv\, du = \tfrac{1}{4}.$$

Podemos obtener la fórmula para cambio de variable en las integrales dobles considerando regiones acotadas inferiormente por el plano $w = 0$, superiormente por el plano $w = 1$, y lateralmente por una superficie cilíndrica y restringiendo la clase de las transformaciones a aquellas en que $w = z$. Sea $\mathscr{E}$ un conjunto en R^3 y sea $\mathscr{F}$ la proyección de $\mathscr{E}$ sobre R^2 (figura 3*a*). Sea **f** una transformación de R^3 en R^3 que satisface las condiciones del teorema 6.16 definido por la regla de correspondencia

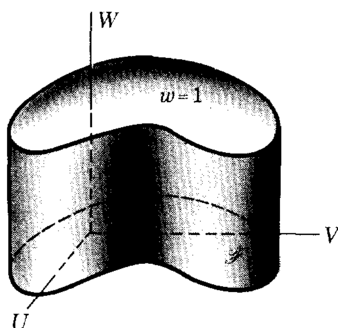
$$\mathbf{y} = \mathbf{f}(\mathbf{x}) = (f_1(u, v), f_2(u, v), w).$$

Entonces, si $\bar{\mathbf{f}}$ es la función de R^2 en R^2 definida por

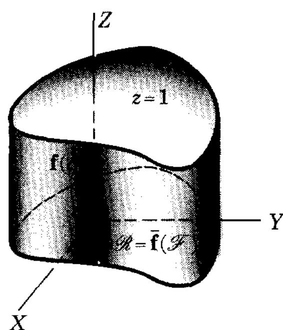
$$\bar{\mathbf{f}}(u, v) = (f_1(u, v), f_2(u, v))$$

tenemos

$$J\mathbf{f}(\mathbf{x}) = \begin{vmatrix} D_1 f_1(u,v) & D_2 f_1(u,v) & 0 \\ D_1 f_2(u,v) & D_2 f_2(u,v) & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} D_1 f_1(u,v) & D_2 f_1(u,v) \\ D_1 f_2(u,v) & D_2 f_2(u,v) \end{vmatrix} = J\bar{\mathbf{f}}(u, v)$$



(a)



(b)

FIGURA 3

Sea $\bar{g}(x, y) = g(x, y, z)$ independiente de z, y que satisfaga las condiciones del teorema 6.16 y sea $\mathscr{R} = \bar{\mathbf{f}}(\mathscr{F})$. Tenemos entonces

$$\iint_{\mathscr{R}} \bar{g}(x, y)\, dx\, dy = \int_0^1 \left[\iint_{\bar{\mathbf{f}}(\mathscr{F})} \bar{g}(x, y)\, dx\, dy \right] dz = \int_{\mathbf{f}(\mathscr{E})} g(\mathbf{y})\, d\mathbf{y}$$

$$= \int_{\mathscr{E}} g(\bar{\mathbf{f}}(\mathbf{x}))\, |J\mathbf{f}(\mathbf{x})|\, d\mathbf{x} = \int_0^1 \left[\iint_{\mathscr{F}} \bar{g}(\bar{\mathbf{f}}(u, v))\, |J\bar{\mathbf{f}}(u, v)|\, du\, dv \right] dw$$

$$= \iint_{\mathscr{F}} \bar{g}(\bar{\mathbf{f}}(u, v))\, |J\mathbf{f}(u, v)|\, du\, dv\,.$$

Luego, si prescindimos de las barras sobre las funciones $\bar{\mathbf{f}}$ y $\bar{g}$, tenemos el siguiente corolario al teorema 6.16.

6.18 Corolario. *Sea* $\mathbf{f}$ *una transformación de* R^2 *en* R^2 *de clase* C^1 *y univalente sobre un conjunto abierto* $\mathscr{G} \subset \mathrm{R}^2$ *con jacobiano* $J\mathbf{f}(u, v) \neq 0$ *para* $(u, v) \in \mathscr{G}$ *cualesquiera y sea* g *una función de* R^2 *en* R *continua sobre* $\mathbf{f}(\mathscr{G})$. *Si* $[\mathbf{a}, \mathbf{b}] \subset \mathscr{G}$ *y* $\mathscr{F} \subset [\mathbf{a}, \mathbf{b}]$ *tiene área, entonces*

$$\int_{\mathscr{R}} g = \int_{\mathscr{R}} g(x, y)\, dx\, dy = \int_{\mathscr{F}} (g \circ \mathbf{f})\, |J\mathbf{f}| = \int_{\mathscr{F}} g(\mathbf{f}(u, v))\, |J\mathbf{f}(u, v)|\, du\, dv$$

donde $\mathscr{R} = \mathbf{f}(\mathscr{F})$.

Nota. Si $J\mathbf{f}(u, v) = 0$ sobre un conjunto de área cero y si $\mathscr{F}$ menos tal conjunto es la unión de un número finito de conjuntos que tienen área cero, entonces como en la nota que siguió al teorema 6.16, puede demostrarse que la fórmula del corolario 6.18 aún es válida.

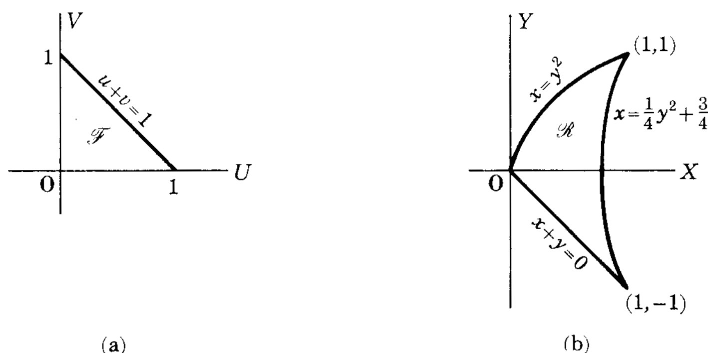


FIGURA 4

6.19 Ejemplo. Sea $\mathbf{f}$ la transformación definida por la regla de correspondencia

$$(x, y) = \mathbf{f}(u, v) = (u^2 + v,\ u - v)$$

y sea $\mathscr{F}$ el triángulo limitado por los ejes coordenados y la recta $u + v = 1$. Encuéntrese el área de $\mathscr{R} = \mathbf{f}(\mathscr{F})$.

SOLUCIÓN. La región $\mathscr{F}$ se muestra en la figura 4*a* y $\mathscr{R} = \mathbf{f}(\mathscr{F})$ se muestra en la figura 4*b*. La transformación $\mathbf{f}$ es univalente para $u \geqslant -\frac{1}{2}$ y para tal caso se tiene:

$$\mathbf{f}: \begin{cases} x = u^2 + v \\ y = u - v \end{cases} \qquad \mathbf{f}^*: \begin{cases} u = -\frac{1}{2} + \sqrt{\frac{1}{4} + x + y} \\ v = -y - \frac{1}{2} + \sqrt{\frac{1}{4} + x + y}. \end{cases}$$

La transformación $\mathbf{f}$ transforma $u = 0$ en $x+y = 0$; $v = 0$ en $x = y^2$; y $u+v = 1$ en $x = \frac{1}{4}y^2+\frac{3}{4}$. Tenemos

$$J\mathbf{f}(u, v) = \begin{vmatrix} 2u & 1 \\ 1 & -1 \end{vmatrix} = -2u-1$$

y

$$A(\mathscr{R}) = \int_0^1 \int_0^{1-u} (2u+1)\, dv\, du = \tfrac{5}{6}.$$

Problemas

1. Evalúese $\int_0^{1/2} \int_y^{1-y} \sqrt{x^2-y^2}\, dx\, dy$ por medio de la transformación $(x, y) = (u-uv, uv)$.

2. Evalúese $\int_0^1 \int_0^x \sqrt{x^2+y^2}\, dy\, dx$ por medio de la transformación $(x, y) = (u, uv)$.

3. Evalúese $\iint_{\mathscr{R}} (x^2+y^2)\, dx\, dy$ donde $\mathscr{R}$ es el paralelogramo con vértices $(1, 1)$, $(5, 2)$, $(6, 5)$ y $(2, 4)$ por medio de la transformación $(x, y) = (1, 1)+u(4, 1)+v(1, 3)$.

4. Un conjunto de puntos de la forma

$$\mathscr{P} = \{P_0+u\mathbf{a}+v\mathbf{b} \mid u, v\in[0, 1]\}$$

es un paralelogramo con vértices $\mathbf{P}_0$, $\mathbf{P}_0+\mathbf{a}$, $\mathbf{P}_0+\mathbf{a}+\mathbf{b}$, y $\mathbf{P}_0+\mathbf{b}$. La transformación $\mathbf{P} = \mathbf{f}(u, v) = \mathbf{P}_0+u\mathbf{a}+v\mathbf{b}$ transforma el cuadrado unitario $\mathscr{S} = \{(u, v) \mid u, v\in[0, 1]\}$ sobre $\mathscr{P}$.

a) Usando la transformación $\mathbf{f}$, encuéntrese el área de $\mathscr{P}$.
b) Encuéntrese el momento de inercia con respecto al eje X del paralelogramo $\mathscr{P} = \{(1, 1)+u(2, 3)+v(1, 5) \mid u, v\in[0, 1]\}$.
c) Encuéntrese el área del paralelogramo con vértices en $(1, 1)$, $(5, 2)$, $(6, 5)$ y $(2, 4)$.

5. *a*) Pruébese que las coordenadas parabólicas (u, v) donde $(x, y) = \mathbf{f}(u, v) = (uv, \frac{1}{2}(u^2-v^2))$ transforman las rectas $u =$ constante y $v =$ constante en parábolas con vértices en $\left(0, \frac{u^2}{2}\right)$ y $\left(0, \frac{v^2}{2}\right)$, respectivamente, y con focos en el origen.
b) Encuéntrese el área de la región $\mathscr{R}$ limitada superiormente por la

parábola $x^2+2y = 1$, inferiormente por la parábola $x^2-2y = 1$, y a la izquierda por el eje Y.

c) Transfórmese la integral de la parte *b* a coordenadas parabólicas y encuéntrese el área de $\mathscr{R}$.

d) Encuéntrese el momento de inercia de $\mathscr{R}$ con respecto al eje Y usando coordenadas parabólicas.

6. *a*) Demuéstrese que las coordenadas elípticas (u, v) donde

$$(x, y) = \mathbf{f}(u, v) = (a \cosh u \cos v, a \operatorname{senh} u \operatorname{sen} v)$$

transforman las rectas u = constante en elipses con focos en $(\pm a, 0)$ y las rectas v = constante en hipérbolas con focos en $(\pm a, 0)$.

b) Encuéntrese el área de la región $\mathscr{R}$ limitada superiormente por la elipse $\dfrac{x^2}{a^2 \cosh^2 1} + \dfrac{y^2}{a^2 \cosh^2 1} = 1$, inferiormente por el eje X y a la izquierda por el eje Y.

c) Pruébese que $J\mathbf{f}(u,v) = a^2(\operatorname{senh}^2 u + \operatorname{sen}^2 v) = \dfrac{a^2}{2}(\cosh 2u - \cos 2v)$

d) Encuéntrese el área de la región $\mathscr{R}$ de la parte *b* usando coordenadas elípticas.

7. Las coordenadas esféricas alargadas (o elípticas) están relacionadas con las coordenadas rectangulares por $(x, y, z) = \mathbf{f}(u, v, w)$ donde

$$\begin{aligned} x &= a\sqrt{(u^2-1)(1-v^2)} \cos w & u &\in [1, \infty\rangle \\ y &= a\sqrt{(u^2-1)(1-v^2)} \operatorname{sen} w & v &\in [-1, 1] \\ z &= auv & w &\in [0, 2\pi] . \end{aligned}$$

a) Demuéstrese que las superficies $u = u_0$ = constante, son los esferoides alargados (elipsoides)

$$\frac{z^2}{{u_0}^2} + \frac{x^2+y^2}{{u_0}^2-1} = a^2$$

y las superficies $v = v_0$ = constante, son los hiperboloides de revolución de dos hojas

$$\frac{z^2}{{v_0}^2} - \frac{x^2+y^2}{1-{v_0}^2} = a^2$$

b) Encuéntrese el jacobiano $J\mathbf{f}(u, v, w)$.

c) Encuéntrese el volumen del esferoide $\frac{z^2}{4} + \frac{x^2}{3} + \frac{y^2}{3} = a^2 (u_0 = 2)$ usando coordenadas rectangulares.

d) Encuéntrese el volumen del esferoide de la parte *c* usando coordenadas esféricas alargadas (*Sugerencia:* en el espacio UVW la superficie limita a u superiormente por 2; v y w toman el rango completo de valores.)

e) Encuéntrese el momento de inercia con respecto al plano XY del esferoide de la parte *c*.

7. COORDENADAS POLARES

La transformación **f** que expresa un punto en el plano R^2 en términos de coordenadas polares viene dada por la regla de correspondencia

$$(x, y) = \mathbf{f}(r, \theta) = (r \cos \theta, r \operatorname{sen} \theta).$$

El jacobiano de **f** es

$$J\mathbf{f}(r,\theta) = \begin{vmatrix} \cos\theta & -r \operatorname{sen}\theta \\ \operatorname{sen}\theta & r\cos\theta \end{vmatrix} = r(\cos^2\theta + \operatorname{sen}^2\theta) = r$$

de modo que al transformar una integral doble en coordenadas polares reemplazamos $dx\, dy$ por $r\, dr\, d\theta$. Las condiciones del corolario 6.18 no están satisfechas por esta transformación. No es univalente y, además, $J\mathbf{f}(\mathbf{x}) = 0$ sobre la recta $r = 0$ en el plano $R\Theta$. Sin embargo, podemos probar que si restringimos el dominio de **f** a una franja

$$\mathscr{S} = \{(r, \theta) \mid 0 \leqslant r, \ \alpha \leqslant \theta \leqslant \alpha + 2\pi\}$$

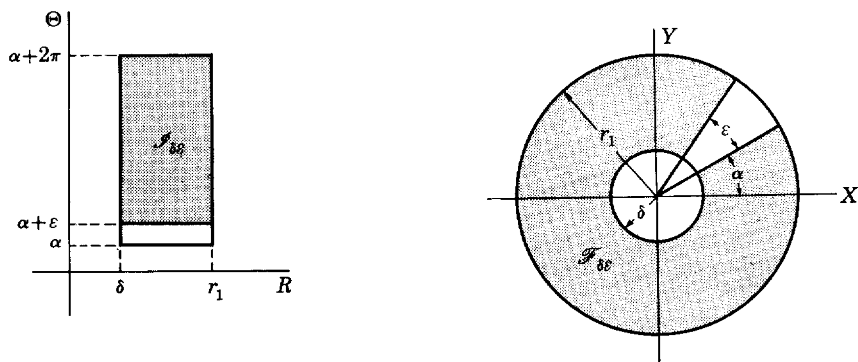


FIGURA 5

en el plano $R\Theta$, entonces la fórmula del corolario 6.18 se verifica para cualquier subconjunto de $\mathscr{S}$ que tenga área. Para ello basta demostrar que se verifica sobre cualquier rectángulo de la forma

$$\mathscr{I} = \{(r, \theta) \mid 0 \leqslant r \leqslant r_1,\ \alpha \leqslant \theta \leqslant \alpha + 2\pi\}.$$

Para cada par de números positivos δ, ε con $\delta < r_1$ y $\varepsilon < 2\pi$, sea

$$\mathscr{I}_{\delta\varepsilon} = \{(r, \theta) \mid \delta \leqslant r \leqslant r_1,\ \alpha + \varepsilon \leqslant \theta \leqslant \alpha + 2\pi\}.$$

La transformación $\mathbf{f}$ es univalente sobre un conjunto abierto que contiene $\mathscr{I}_{\delta\varepsilon}$ y $J\mathbf{f}(r, \theta) = r \neq 0$ sobre $\mathscr{I}_{\delta\varepsilon}$ La transformación $\mathbf{f}$ transforma $\mathscr{I}_{\delta\varepsilon}$ sobre $\mathbf{f}(\mathscr{I}_{\delta\varepsilon}) = \mathscr{F}_{\delta\varepsilon}$ donde $\mathscr{F}_{\delta\varepsilon}$ es el anillo limitado por las circunferencias de radios δ y r_1 con una muesca de ángulo ε quitada (figura 5). De aquí que si g es continua sobre $\mathbf{f}(\mathscr{I})$, la fórmula del corolario 6.18 se verifica sobre $\mathscr{I}_{\delta\varepsilon}$ e

$$\iint_{\mathscr{F}_{\delta\varepsilon}} g(x, y)\, dx\, dy = \iint_{\mathscr{I}_{\delta\varepsilon}} g(r \cos \theta,\ r \operatorname{sen} \theta) r\, dr\, d\theta$$

para δ y ε positivos cualesquiera. Tomando el límite cuando δ y ε tienden a cero obtenemos

$$\iint_{\mathscr{F}} g(x, y)\, dx\, dy = \iint_{\mathscr{I}} g(r \cos \theta,\ r \operatorname{sen} \theta) r\, dr\, d\theta$$

donde $\mathscr{F}$ es el disco de radio r_1 con centro en el origen. De donde si $\mathscr{R}$ es cualquier conjunto en $\mathscr{S}$ que tiene área y si $\mathscr{F} = \mathbf{f}(\mathscr{R})$, entonces

7.1 $$\iint_{\mathscr{F}} g(x, y)\, dx\, dy = \iint_{\mathscr{R}} g(r \cos \theta,\ r \operatorname{sen} \theta) r\, dr\, d\theta.$$

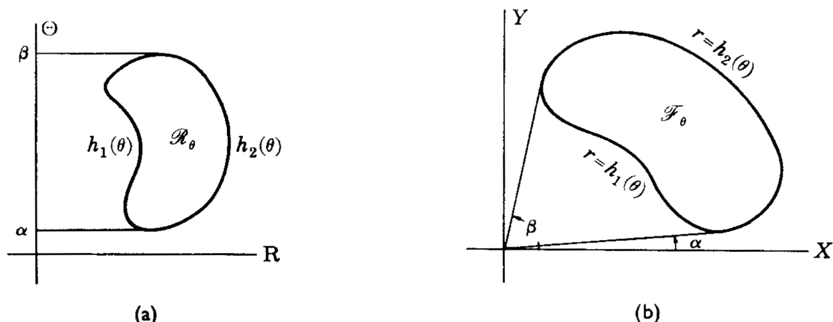


FIGURA 6

Consideremos un conjunto de la forma $\mathscr{R}_0 = \{(r, \theta) \mid \alpha \leqslant \theta \leqslant \beta, 0 \leqslant h_1(\theta) \leqslant r \leqslant h_2(\theta)\}$, donde $0 \leqslant \beta - \alpha \leqslant 2\pi$ y h_1, h_2 son funciones continuas sobre el intervalo $[\alpha, \beta]$ con $0 \leqslant h_1(\theta) \leqslant h_2(\theta)$ para toda $\theta \in [\alpha, \beta]$. La región correspondiente en el plano XY es $\mathscr{F}_0 = \{r \cos\theta, r \operatorname{sen}\theta) \mid \alpha \leqslant \theta \leqslant \beta, 0 \leqslant h_1(\theta) \leqslant r \leqslant h_2(\theta)\}$ (figura 6). Si g es continua sobre $\mathscr{F}_0$, entonces

$$\int_{\mathscr{F}_\theta} g = \iint_{\mathscr{R}_\theta} g(r\cos\theta,\ r\operatorname{sen}\theta)\, r\, dr\, d\theta = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} g(r\cos\theta,\ r\operatorname{sen}\theta)\, r\, dr\, d\theta .$$

Si el integrando es la función constante 1, entonces $\int_{\mathscr{F}_\theta} 1$ es el área de $\mathscr{F}_\theta$ y

$$A(\mathscr{F}_\theta) = \int_{\mathscr{F}_\theta} 1 = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} r\, dr\, d\theta .$$

7.2 Ejemplo. Encuéntrese el área de la región que se localiza en el interior del círculo $r = 3 \operatorname{sen}\theta$ y en el exterior de la cardioide $r = 1 + \operatorname{sen}\theta$.

SOLUCIÓN. El área deseada es la de la región sombreada en la figura 7. Se encuentra el límite de θ haciendo

$$3 \operatorname{sen}\theta = 1 + \operatorname{sen}\theta$$

y resolviendo:

$$\operatorname{sen}\theta = \frac{1}{2}, \qquad \theta = \frac{\pi}{6}, \frac{5\pi}{6} .$$

Entonces

$$A(\mathscr{F}_\theta) = \int_{\mathscr{F}\theta} 1 = \int_{\pi/6}^{5\pi/6} \int_{1+\operatorname{sen}\theta}^{3\operatorname{sen}\theta} r\, dr\, d\theta = \pi .$$

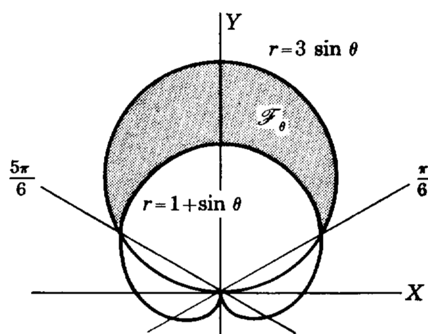


FIGURA 7

7.3 Ejemplo. Encuéntrese el centroide de la región del ejemplo 7.2.

SOLUCIÓN. Como la región $\mathscr{F}_\theta$ es simétrica respecto al rayo $\theta = \frac{\pi}{2}$, el centroide se encuentra sobre este rayo. Necesitamos, pues, solamente encontrar el momento con respecto al eje X y dividir por el área para encontrar la segunda coordenada del centroide.

$$M_x = \int_{\mathscr{F}_\theta} y\,d\mathbf{x} = \int_{\pi/6}^{5\pi/6} \int_{1+\operatorname{sen}\theta}^{3\operatorname{sen}\theta} r^2 \operatorname{sen}\theta\,dr\,d\theta = \frac{11\pi}{6} + \frac{9\sqrt{3}}{16}.$$

Así pues

$$\bar{y} = \frac{11}{6} + \frac{9\sqrt{3}}{16\pi} \approx 2.14$$

de modo que las coordenadas rectangulares del centroide son $\left(0,\ \frac{11}{6} + \frac{9\sqrt{3}}{16\pi}\right)$ $\approx (0{,}2.14)$ y sus coordenadas polares $\left(\frac{11}{6} + \frac{9\sqrt{3}}{16\pi},\ \frac{\pi}{2}\right) \approx \left(2.14,\ \frac{\pi}{2}\right)$.

Nuestra discusión previa (pág. 365) del volumen de una región en R^3 limitada lateralmente por una superficie cilíndrica con un generador paralelo al eje Z, superiormente limitada por una superficie $z = g(x, y)$ e inferiormente por una región cerrada y acotada en el plano XY se sigue aplicando en el caso en que la región está descrita en coordenadas cilíndricas

$$(x, y, z) = (r\cos\theta,\ r\operatorname{sen}\theta,\ z).$$

Si una tal región de R^3 tiene $\mathscr{F}_\theta$ como su frontera inferior, entonces

7.4
$$V = \int_{\mathscr{F}_\theta} g = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} g(r\cos\theta,\ r\operatorname{sen}\theta)\,r\,dr\,d\theta.$$

7.5 Ejemplo. Encuéntrese el volumen de la región limitada superiormente por el paraboloide de revolución $z = x^2+y^2$ inferiormente por el plano XY, y lateralmente por el cilindro circular $x^2+y^2 = 4$.

SOLUCIÓN. (Figura 20, pág. 367.) En coordenadas cilíndricas, la región está limitada superiormente por la superficie $z = r^2$ y lateralmente por el cilindro circular $r = 2$. La región base $\mathscr{F}_\theta = \{(r\cos\theta,\ r\operatorname{sen}\theta) \mid 0 \leqslant r \leqslant 2, 0 \leqslant \theta \leqslant 2\pi\}$. De donde

$$V = \int_0^{2\pi} \int_0^2 r^2\,r\,dr\,d\theta = \int_0^{2\pi} \int_0^2 r^3\,dr\,d\theta = 8\pi.$$

Problemas

1. Encuéntrese el área de cada una de las siguientes regiones.

a) Limitada por la cardioide $r = a(1 + \cos\theta)$
b) dentro de la cardioide $r = a(1 + \cos\theta)$ y fuera del círculo $r = a$
c) interior del círculo $r = 6$ y a la derecha de la recta $r\cos\theta = 3$
d) interior del círculo $r = \cos\theta + \operatorname{sen}\theta$ y exterior del círculo $r = 1$
e) interior del círculo $r = \cos\theta$ y exterior de la cardioide $r = 1 - \cos\theta$
f) limitada por una de las hojas de la rosa de cuatro hojas $r = a\operatorname{sen} 2\theta$
g) limitada por el bifolio $r = a\operatorname{sen}\theta\cos^2\theta$
h) limitada por una hoja de la lemniscata $r^2 = 2a^2\cos 2\theta$.

2. Encuéntrese el centroide de cada una de las siguientes regiones.

a) Problema 1*a*
b) Problema 1*c*
c) Problema 1*e*
d) Problema 1*g*
e) Problema 1*d*.
(*Sugerencia.* $M_x = M_y$ por simetría.)
f) Limitada por la parábola $r = \dfrac{d}{1 - \cos\theta}$ y el eje Y
g) limitada por una hoja de la lemniscata $r^2 = 2a^2\cos 2\theta$.
(*Sugerencia.* Hágase $\sqrt{\cos 2\theta} = \operatorname{sen}\varphi$.)
h) Limitada por el lazo de la estrofoide $r = a\cos 2\theta\sec\theta$.

3. Encuéntrese el momento de inercia de cada una de las siguientes regiones con respecto a la recta que se indica.

a) limitada por una sola hoja de la rosa de cuatro hojas $r = a\operatorname{sen} 2\theta$ con respecto al eje Y. (*Sugerencia:* adviértase que por simetría $I_x = I_y$ y que $I_x + I_y$ es más fácil calcular que I_y)
b) interior del círculo $r = \cos\theta + \operatorname{sen}\theta$ y exterior del círculo $r = 1$ con respecto al eje X
c) limitada por una hoja de la lemniscata $r^2 = 2a^2\cos 2\theta$ con respecto al eje Y
d) interior de la cardioide $r = a(1 + \cos\theta)$ y exterior del círculo $r = a$ con respecto al eje Y.

4. Encuéntrese el volumen de una semiesfera de radio la unidad usando coordenadas cilíndricas.

5. Encuéntrese el volumen de la región de R^3 limitada arriba por $z = 1 - x^2 - y^2$ y abajo por el plano XY usando coordenadas cilíndricas.

6. Encuéntrese el volumen de la región de R^3 limitada arriba por $z = 4 - x$, lateralmente por $x^2 + y^2 = 1$, y debajo por el plano XY, usando coordenadas cilíndricas.

7. Encuéntrese el volumen de la región del primer octante limitada arriba por $z = x$ y lateralmente en el interior por $r = 1 + \operatorname{sen}\theta$ y en el exterior por $r = 3\operatorname{sen}\theta$.

8. Demuéstrese que la región de R^3 que se encuentra sobre

$$\mathscr{F}_\theta = \{(r\cos\theta, r\operatorname{sen}\theta) \mid 0 \leqslant \theta \leqslant \alpha, r_1 \leqslant r \leqslant r_2\}$$

y está limitada superiormente por el plano $z = h$ tiene volumen $\alpha\bar{r}(r_2 - r_1)h$ donde $\bar{r} = \frac{1}{2}(r_2 + r_1)$.

9. *a*) Encuéntrese el volumen del segmento esférico de una base limitado por la esfera $r^2 + z^2 = a^2$ y el plano $z = a - h$ donde $0 < h < a$.

b) Encuéntrese el volumen del segmento esférico de dos bases limitado por la esfera $r^2 + z^2 = a^2$ y los planos $z = a - h_1$ y $z = a - h_2$ donde $0 \leqslant h_1 < h_2 < a$.

8. COORDENADAS ESFÉRICAS

La transformación **f** que expresa un punto de R^3 en términos de las coordenadas esféricas viene dada por la regla de correspondencia

$$(x, y, z) = \mathbf{f}(\rho, \theta, \varphi) = (\rho\operatorname{sen}\varphi\cos\theta, \rho\operatorname{sen}\varphi\operatorname{sen}\theta, \rho\cos\varphi)$$

donde $\rho \geqslant 0$, $0 \leqslant \theta \leqslant 2\pi$, $0 \leqslant \varphi \leqslant \pi$. El jacobiano de **f** es

$$J\mathbf{f}(\rho, \theta, \varphi) = \begin{vmatrix} \operatorname{sen}\varphi\cos\theta & -\rho\operatorname{sen}\varphi\operatorname{sen}\theta & \rho\cos\varphi\cos\theta \\ \operatorname{sen}\varphi\operatorname{sen}\theta & \rho\operatorname{sen}\varphi\cos\theta & \rho\cos\varphi\operatorname{sen}\theta \\ \cos\varphi & 0 & -\rho\operatorname{sen}\varphi \end{vmatrix}$$
$$= -\rho^2\operatorname{sen}\varphi$$

de modo que al transformar una integral triple a coordenadas esféricas reemplazamos $dx\,dy\,dz$ por $|J\mathbf{f}(\rho, \theta, \varphi)|\,d\rho\,d\theta\,d\varphi$. Aquí las condiciones del teorema 6.16 no son satisfechas. Tenemos $J\mathbf{f}(\rho, \theta, \varphi) = 0$ si $\rho = 0$, $\varphi = 0$ o $\varphi = \pi$ y **f** no es univalente sobre las fronteras del dominio de **f**. Sin embargo, un argumento análogo al usado para las coordenadas polares muestra que la fórmula del teorema 6.16 se verifica también para las coordenadas esféricas y

8.1
$$\iiint_{\mathbf{f}(\mathscr{E})} g(x, y, z)\,dx\,dy\,dz$$
$$= \iiint_{\mathscr{E}} g(\rho\operatorname{sen}\varphi\cos\theta, \rho\operatorname{sen}\varphi\operatorname{sen}\theta, \rho\cos\varphi)\rho^2\operatorname{sen}\varphi\,d\rho\,d\theta\,d\varphi\,.$$

8.2 Ejemplo. Encuéntrese el volumen de la región limitada por una esfera de radio a.

SOLUCIÓN.

$$V = \int_{\mathbf{f}(\mathscr{E})} d\mathbf{y} = \int_0^{\pi} \int_0^{2\pi} \int_0^{a} \rho^2 \operatorname{sen} \varphi \, d\rho \, d\theta \, d\varphi = \frac{4\pi a^3}{3}.$$

8.3 Ejemplo. Encuéntrese el centroide de la región limitada por una semiesfera de radio a.

SOLUCIÓN. Supongamos que el hemisferio es la mitad superior de la esfera del ejemplo 8.2. Por simetría concluimos que el centroide se encuentra sobre el eje Z. Necesitamos solamente calcular, por tanto, $M_{xy} = \int_{\mathbf{f}(\mathscr{E})} z \, d\mathbf{y}$. Como $z = \rho \cos \varphi$, tenemos

$$M_{xy} = \int_{\mathbf{f}(\mathscr{E})} z \, d\mathbf{y} = \int_0^{\pi/2} \int_0^{2\pi} \int_0^{a} \rho \cos \varphi \, \rho^2 \operatorname{sen} \varphi \, d\rho \, d\theta \, d\varphi = \frac{\pi a^4}{4}.$$

Por tanto

$$\bar{z} = \frac{M_{xy}}{V} = \frac{\pi a^4}{4} \Big/ \frac{2\pi a^3}{3} = \frac{3a}{8} \quad \text{y} \quad \bar{\mathbf{x}} = (\bar{x}, \bar{y}, \bar{z}) = \left(0, 0, \frac{3a}{8}\right).$$

Problemas

1. Úsese la integración en coordenadas esféricas para encontrar el volumen de la región acotada por la esfera $x^2+y^2+z^2 = 4$ superiormente y por el cono $z^2 = 3(x^2+y^2)$ inferiormente.

2. Encuéntrese el centroide de la región del problema 1.

3. Encuéntrese el momento de inercia de la región del problema 1 respecto al eje Z y úsese esto para encontrar los momentos de inercia con respecto al plano XZ y al plano YZ.

4. Encuéntrese el momento de inercia de la región limitada por una esfera de radio a con respecto a un diámetro.

5. Úsense coordenadas esféricas para encontrar el volumen del esferoide $b^2(x^2+y^2)+a^2z^2 = a^2b^2$. (*Sugerencia:* después de integrar con respecto a ρ y θ, hágase $\cos \varphi = u$.)

6. Encuéntrese el centroide de la región en el primer octante limitada por la esfera $\rho = a$ y los planos coordenados. (*Sugerencia:* encuéntrese $\bar{z}$ y úsese la simetría.)





Sucesiones

1. INTRODUCCIÓN

Una sucesión es una función que tiene como dominio el conjunto de los enteros positivos. Así pues, una sucesión es una función de una variable real. En el capítulo 3 se consideraron las funciones vectoriales de una variable real, pero la mayor parte del análisis de tales funciones que allí vimos no se aplica a las sucesiones. Las operaciones de diferenciación e integración se definieron para funciones cuyo dominio es un intervalo no degenerado o una unión de intervalos no degenerados. (Por intervalo no degenerado entendemos un intervalo que contiene más de un punto.) Tales funciones a veces se llaman funciones de una variable real continua. Pero las sucesiones son funciones definidas solamente sobre un conjunto

discreto de puntos de la recta y, por tanto, las operaciones de diferenciación e integración no se definen para las sucesiones.

La noción de límite de una función de variable real se aplica a las sucesiones y en este capítulo estudiaremos detalladamente el concepto de límite. Recuérdese que el límite de una sucesión se define solamente en puntos de acumulación del dominio de la función. Aquí incluiremos los puntos ideales ∞ y $-\infty$ como posibles puntos de acumulación del dominio de una función. Un punto p es un punto de acumulación de un conjunto si toda vecindad de p contiene un punto del conjunto distinto de p. Si p es un número real, entonces una vecindad de p es un intervalo abierto $\langle a, b\rangle$ que contiene a p. Si p es $\infty(-\infty)$, entonces una vecindad de p es un intervalo $\langle a, \infty\rangle$ ($\langle -\infty, b\rangle$). Como el dominio de una sucesión es el conjunto de los enteros positivos, ∞ es el único punto de acumulación del dominio de una sucesión. Por tanto, para sucesiones, sólo consideraremos límites en ∞.

Si en el capítulo 3 hubiésemos considerado límites en ∞ de funciones vectoriales de una variable real, entonces el límite de una sucesión sería un caso particular de lo que allí habríamos visto. Como tales límites no fueron considerados entonces, nuestra discusión de límites de sucesiones principia desde el comienzo en este capítulo. Sin embargo, mucho de lo que aquí vamos a hacer es muy semejante a lo que hicimos al estudiar los límites en el capítulo 3.

2. LÍMITE DE UNA SUCESIÓN

2.1 Definición. *Una **sucesión** de puntos en* R^m *es una función cuyo dominio es el conjunto de enteros positivos y cuyo rango es un conjunto de puntos en* R^m.

Así pues, una sucesión de puntos $\mathbf{s}$ es una correspondencia del conjunto de los enteros positivos a un conjunto de puntos; es decir, para cada entero positivo n hay un punto $\mathbf{s}(n)$ que le corresponde. Es más común escribir $\mathbf{s}_n$ que $\mathbf{s}(n)$ y denotar la sucesión por $\{\mathbf{s}_n\}$ en lugar de por $\mathbf{s}$. El punto $\mathbf{s}_n$ se llama *n-ésimo término* de la sucesión.

Un ejemplo de una sucesión $\{\mathbf{s}_n\}$ de puntos en R viene dado por la regla de correspondencia $\mathbf{s}_n = \dfrac{1}{n}$. Esta sucesión también puede describirse escribiendo cierto número de sus primeros términos en orden: $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$

La regla de correspondencia $\mathbf{s}_n = \left(n, n^2, \dfrac{1}{n}\right)$ describe una sucesión de puntos en R^3: $(1, 1, 1), (2, 4, \frac{1}{2}), (3, 9, \frac{1}{3}), \ldots$

Como una sucesión $\{\mathbf{s}_n\}$ de puntos en R^m es una función de R en R^m cuyo dominio tiene ∞ como su único punto de acumulación, la noción de límite

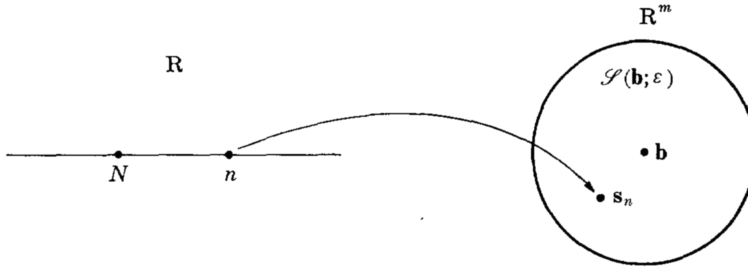


FIGURA 1

sólo puede definirse en ∞. Por tanto, el límite de $\{\mathbf{s}_n\}$ en ∞ puede llamarse simplemente el límite de $\{\mathbf{s}_n\}$ y puede denotarse por lím $\mathbf{s}_n$ lo mismo que por $\lim\limits_{n\to\infty} \mathbf{s}_n$.

2.2 Definición. *El* ***límite*** *de* $\{\mathbf{s}_n\}$ *es* $\mathbf{b}$, *lo que se denota por* lím $\mathbf{s}_n = \mathbf{b}$ *o* $\lim\limits_{n\to\infty} \mathbf{s}_n = \mathbf{b}$, *si para cada* $\varepsilon > 0$ *existe un número* N *tal que* $|\mathbf{s}_n - \mathbf{b}| < \varepsilon$ *siempre que* $n > N$.

Reformulando esta definición en términos de vecindades, tenemos que lím $\mathbf{s}_n = \mathbf{b}$ si para cada vecindad $\mathscr{S}(\mathbf{b};\varepsilon)$ de $\mathbf{b}$ existe una vecindad $\langle N, \infty\rangle$ de ∞ tal que $\mathbf{s}_n \in \mathscr{S}(\mathbf{b};\varepsilon)$ siempre que $n \in \langle N, \infty\rangle$ (figura 1). Intuitivamente, lím $\mathbf{s}_n = \mathbf{b}$ significa que $\mathbf{s}_n$ está "próximo" a $\mathbf{b}$ para todo n suficientemente grande.

Si lím $\mathbf{s}_n$ existe, entonces decimos que la sucesión $\{\mathbf{s}_n\}$ *converge*. Si una sucesión no converge, entonces decimos que *diverge*.[1]

2.3 Ejemplo. Demuéstrese que para $s_n = \dfrac{1}{n}$, lím $s_n = 0$.

SOLUCIÓN. Sea $\varepsilon > 0$ cualquiera. Deseamos demostrar que existe un número N tal que

$$|s_n - 0| = \left|\frac{1}{n} - 0\right| = \frac{1}{n} < \varepsilon \quad \text{siempre que } n > N.$$

Si hacemos $N = \dfrac{1}{\varepsilon}$, entonces $n > N$ implica $\dfrac{1}{n} < \dfrac{1}{N} = \varepsilon$. Así pues, hemos probado que lím $\dfrac{1}{n} = 0$.

[1] En este caso algunos autores dicen que $\{\mathbf{s}_n\}$ "no converge", reservando decir que $\{\mathbf{s}_n\}$ "diverge" cuando $\{\mathbf{s}_n\}$ tiene como límite ∞ o $-\infty$. [N. del T.]

2.4 Ejemplo. Demuéstrese que lím $r^n = 0$ si $0 < |r| < 1$.

SOLUCIÓN. Sea $\varepsilon > 0$ cualquiera. Queremos probar que existe un número N tal que

$$|r^n - 0| = |r|^n < \varepsilon \quad \text{siempre que } n > N.$$

Ahora bien, $|r|^n < \varepsilon$ si $n \ln |r| < \ln \varepsilon$, es decir, si $n > \dfrac{\ln \varepsilon}{\ln |r|}$. Notese que $\ln |r| < 0$ ya que $|r| < 1$. Así pues, si $N = \dfrac{\ln \varepsilon}{\ln |r|}$, entonces $n > N$ implica $|r|^n < \varepsilon$.

2.5 Ejemplo. Demuéstrese que lím$\left(\dfrac{1}{n}, \dfrac{1}{2^n}\right) = (0, 0)$.

SOLUCIÓN. Sea $\varepsilon > 0$. Queremos probar que existe un número N tal que

$$\left|\left(\frac{1}{n}, \frac{1}{2^n}\right) - (0, 0)\right| < \varepsilon \quad \text{siempre que } n > N.$$

Ahora bien

$$\left|\left(\frac{1}{n}, \frac{1}{2^n}\right) - (0, 0)\right| = \sqrt{\frac{1}{n^2} + \frac{1}{2^{2n}}};$$

luego

$$\left|\left(\frac{1}{n}, \frac{1}{2^n}\right) - (0, 0)\right| < \varepsilon \text{ si } \frac{1}{n} < \frac{\varepsilon}{\sqrt{2}} \quad \text{y} \quad \frac{1}{2^n} < \frac{\varepsilon}{\sqrt{2}}.$$

Como lím $\dfrac{1}{n} = 0$, existe un número N_1 $\left(\text{por ejemplo, } N_1 = \dfrac{\sqrt{2}}{\varepsilon}\right)$ tal que $\dfrac{1}{n} < \dfrac{\varepsilon}{\sqrt{2}}$ siempre que $n > N_1$. Por otra parte, lím $\dfrac{1}{2^n} = 0$ implica que existe un número N_2 $\left(\text{por ejemplo, } N_2 = \dfrac{1}{2} - \dfrac{\ln \varepsilon}{\ln 2}\right)$ tal que $\dfrac{1}{2^n} < \dfrac{\varepsilon}{\sqrt{2}}$ siempre que $N > N_2$. Entonces, si $N = \text{máx}\{N_1, N_2\}$

$$\left|\left(\frac{1}{n}, \frac{1}{2^n}\right) - (0, 0)\right| = \sqrt{\frac{1}{n^2} + \frac{1}{2^{2n}}} < \sqrt{\frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2}} = \varepsilon$$

siempre que $n > N$. Con lo que hemos demostrado que lím $\left(\dfrac{1}{n}, \dfrac{1}{2^n}\right) = (0, 0)$.

Nótese que en este ejemplo el límite de la sucesión de puntos es un punto cuyos componentes son los límites de las sucesiones componentes. Esto es cierto en todos los casos.

2.6 Teorema. *Sea* $\mathbf{b} = (b_1, \ldots, b_m)$ *y* $\mathbf{s}_n = (s_n^{\,1}, \ldots, s_n^{\,m})$, *donde* $s_n^{\,k}$ $(k = 1, \ldots, m)$ *es el k-ésimo componente de* $\mathbf{s}_n$. *Entonces* lím $\mathbf{s}_n = \mathbf{b}$ *si y sólo si* $s_n^{\,k} = b_k$ *para cada* $k = 1, \ldots, m$.

No se da la prueba de este teorema por ser esencialmente la misma que la prueba del teorema 3.3, pág. 00. Una aplicación importante de este teorema es la determinación del límite de una sucesión de números complejos. Los números complejos $z_n = x_n + iy_n$ pueden considerarse como puntos (x_n, y_n) de R^2. Entonces

$$\text{lím } z_n = (\text{lím } x_n,\ \text{lím } y_n) = \text{lím } x_n + i \text{ lím } y_n$$

si lím x_n y lím y_n existen. Así por ejemplo,

$$\text{lím}\left(\frac{1}{n} + \frac{i}{2^n}\right) = 0 + 0i = 0.$$

Si formamos una sucesión prescindiendo de algunos de los términos de una sucesión dada, entonces esta nueva sucesión se llama subsucesión de la sucesión original.

2.7 Definición. *Si* $\{\mathbf{s}_n\}$ *es una sucesión y* $\{n_k\}$ *es una sucesión creciente de enteros positivos, entonces la sucesión* $\{\mathbf{s}_{n_k}\}$ *se llama* ***subsucesión*** *de la* $\{\mathbf{s}_n\}$.

Nótese que $\mathbf{s}_{n_k}$ está definida mediante una composición de funciones. Esto se hace más aparente si escribimos $\mathbf{s}_{n_k}$ en la forma $\mathbf{s}(n(k))$. Por ejemplo, si $s_n = \dfrac{1}{n}$ y $n_k = 2k$, entonces $s_{n_k} = s(n(k)) = s(2k) = \dfrac{1}{2k}$ y, por tanto, $\left\{\dfrac{1}{2k}\right\}$ es una subsucesión de $\left\{\dfrac{1}{n}\right\}$. Es la subsucesión consistente en todos los términos pares de la sucesión $\left\{\dfrac{1}{n}\right\}$.

2.8 Teorema. *Si* $\{\mathbf{s}_n\}$ *converge, entonces cualquier subsucesión de la* $\{\mathbf{s}_n\}$ *converge al mismo punto.*

PRUEBA. Si $\lim\limits_{n\to\infty} \mathbf{s}_n = \mathbf{b}$, deseamos demostrar que, para cualquier subsucesión $\{\mathbf{s}_{n_k}\}$ de $\{\mathbf{s}_n\}$, $\lim\limits_{k\to\infty} \mathbf{s}_{n_k} = \mathbf{b}$. Es decir, queremos probar que para cualquier $\varepsilon > 0$ existe un número K tal que $\mathbf{s}_{n_k} \in \mathscr{S}(\mathbf{b}; \varepsilon)$ para todo $k > K$. Tomemos $\varepsilon > 0$. Como lím $\mathbf{s}_n = \mathbf{b}$, existe un número N tal que $\mathbf{s}_n \in \mathscr{S}(\mathbf{b}; \varepsilon)$

para todo $n > N$. Además, como $\{n_k\}$ es una sucesión creciente de números positivos existe un número K tal que $n_k > N$ siempre que $k > K$. Así pues, si $k > K$, $n_k > N$ y, por tanto, $\mathbf{s}_{n_k} \in \mathscr{S}(\mathbf{b}; \varepsilon)$.

2.9 Ejemplo. Pruébese que $\left\{\dfrac{1}{n+p}\right\}$, donde p es algún entero positivo, puede considerarse como una subsucesión de $\left\{\dfrac{1}{n}\right\}$ y, por tanto, $\lim\limits_{n\to\infty} \dfrac{1}{n+p} = 0$.

SOLUCIÓN. Si $s_n = \dfrac{1}{n}$ y $n_k = k+p$, entonces $s_{n_k} = \dfrac{1}{k+p}$ y $\left\{\dfrac{1}{k+p}\right\}$ es una subsucesión de $\left\{\dfrac{1}{n}\right\}$. Por tanto, por el teorema 2.8, $\lim\limits_{k\to\infty} \dfrac{1}{k+p} = \lim\limits_{n\to\infty} \dfrac{1}{n} = 0$.

Problemas

1. Escríbanse los primeros cinco términos de cada una de las siguientes sucesiones $\{s_n\}$:

a) $s_n = n^2$ *b*) $s_n = c$ *c*) $s_n = (-1)^n$

d) $s_n = \dfrac{1}{n^3}$ *e*) $s_n = 1 + \dfrac{1}{n}$ *f*) $s_n = 1 + (-1)^n \dfrac{1}{n}$.

2. Usando la definición 2.2, verifíquese lo siguiente:

a) $\lim c = c$ *b*) $\lim \dfrac{1}{n^2} = 0$

c) $\lim \dfrac{1}{n^3} = 0$ *d*) $\lim \left(1 + \dfrac{1}{n}\right) = 1$

e) $\lim \left(1 + (-1)^n \dfrac{1}{n}\right) = 1$ *f*) $\lim \dfrac{1}{4n} = 0$.

3. Úsese la definición 2.2 para probar que $\lim\limits_{n\to\infty} \dfrac{1}{np} = 0$, donde p es algún entero positivo.

4. ¿Convergen las siguientes sucesiones de puntos? Proporciónense los límites de las que convergen.

a) $\mathbf{s}_n = \left(1, \dfrac{1}{n}\right)$ *b*) $\mathbf{s}_n = \left(\dfrac{1}{2^n}, \dfrac{1}{n^2}, \dfrac{1}{5^n}\right)$

c) $\mathbf{s}_n = \left(n^2, \dfrac{1}{n^3}\right)$ *d*) $\mathbf{s}_n = \left(5, 1 + \dfrac{1}{n}, \dfrac{1}{(-2)^n}\right)$.

5. Proporciónense los límites de las siguientes funciones complejas:

a) $\left\{3 + \dfrac{i}{n^2}\right\}$ *b*) $\left\{\dfrac{1}{2^n} + i\right\}$.

6. Pruébese que $\lim \mathbf{s}_n = \mathbf{0}$ si y sólo si $\lim |\mathbf{s}_n| = 0$.

7. Si $\lim\limits_{k\to\infty} \mathbf{s}_{p+k} = \mathbf{b}$ donde p es algún entero positivo, demuéstrese que $\lim\limits_{n\to\infty} \mathbf{s}_n = \mathbf{b}$.

8. Demuéstrese que $\left\{\dfrac{1}{pn}\right\}$ y $\left\{\dfrac{1}{n^p}\right\}$, donde p es un cierto entero cualquiera positivo, pueden ambas considerarse como subsucesiones de $\left\{\dfrac{1}{n}\right\}$ y, por tanto, deben convergir a 0.

9. Si las sucesiones $\{x_n\}$ y $\{y_n\}$ convergen y $x_n \leqslant y_n$ para todo n, pruébese que $\lim x_n \leqslant \lim y_n$. ¿Cuál es la conclusión si $x_n \leqslant y_n$ para toda n?

10. Si $\{\mathbf{s}_n\}$ es una sucesión de puntos tales que $\lim\limits_{k\to\infty} \mathbf{s}_{2k} = \mathbf{b}$ y $\lim\limits_{n\to\infty} \mathbf{s}_{2k+1} = \mathbf{b}$, muéstrese que $\lim\limits_{n\to\infty} \mathbf{s}_n = \mathbf{b}$.

3. CONVERGENCIA DE SUCESIONES

Según el teorema 2.6 se ve que las cuestiones de convergencia de sucesiones de puntos de R^n pueden reducirse a cuestiones sobre convergencia de sucesiones de números reales. En esta sección probaremos algunos teoremas que son útiles en la determinación de límites de sucesiones específicas de números reales.

3.1 Teorema. *Si* $\mathbf{f}$ *es una función de* R^m *a* R^n *que es continua en el punto* $\mathbf{p}$ *y* $\{\mathbf{s}_n\}$ *es una sucesión de puntos en* $\mathscr{D}_f$ *tales que* $\lim \mathbf{s}_n = \mathbf{p}$, *entonces* $\lim \mathbf{f}(\mathbf{s}_n) = \mathbf{f}(\mathbf{p})$.

PRUEBA. Tomemos $\varepsilon > 0$. Deseamos demostrar que existe un número N tal que $\mathbf{f}(\mathbf{s}_n) \in \mathscr{S}(\mathbf{f}(\mathbf{p}); \varepsilon)$ siempre que $n > N$. Como $\mathbf{f}$ es continua en $\mathbf{p}$, existe una vecindad $\mathscr{S}(\mathbf{p}; \delta)$ de $\mathbf{p}$ tal que $\mathbf{f}(\mathbf{x}) \in \mathscr{S}(\mathbf{f}(\mathbf{p}); \varepsilon)$ siempre que $\mathbf{x} \in \mathscr{S}(\mathbf{p}; \delta) \cap \mathscr{D}_f$. Además, como $\lim \mathbf{s}_n = \mathbf{p}$ y $\mathbf{s}_n \in \mathscr{D}_f$, existe un número N tal que $\mathbf{s}_n \in \mathscr{S}(\mathbf{p}; \delta) \cap \mathscr{D}_f$ siempre que $n > N$. Por tanto, si $n > N$, $\mathbf{f}(\mathbf{s}_n) \in \mathscr{S}(f(\mathbf{p}); \varepsilon)$ y esto completa la prueba.

3.2 Corolario. *Si* lím $x_n = x$ *y* lím $y_n = y$, *entonces*

1) lím $(x_n + y_n) = x + y$
2) lím $(x_n - y_n) = x - y$
3) lím $(x_n y_n) = xy$
4) lím $\dfrac{x_n}{y_n} = \dfrac{x}{y}$, *si* $y_n \neq 0$ *para toda* n *y* $y \neq 0$.

PRUEBA. Solamente probaremos 1). Las pruebas de las restantes afirmaciones son análogas. Sea f la función definida por $f(u, v) = u + v$, $\mathbf{s}_n = (x_n, y_n)$, y $\mathbf{p} = (x, y)$. Entonces lím $\mathbf{s}_n = \mathbf{p}$. Como f es un polinomio, es continuo en todos los puntos en R^2. Por tanto, de acuerdo con el teorema 3.1,

$$\text{lím } (x_n + y_n) = \text{lím } f(\mathbf{s}_n) = f(\mathbf{p}) = x + y.$$

3.3 Ejemplo. Determínese lím $\dfrac{3n^2 - 5n + 2}{n^2 + 7n - 4}$.

SOLUCIÓN. No podemos aplicar el corolario 3.2 a esta sucesión en la forma en que aparece ya que los límites del numerador y del denominador no existen. Sacando como factor en el numerador y en el denominador la máxima potencia con que n aparece en ellos, obtenemos

$$\frac{3n^2 - 5n + 2}{n^2 + 7n - 4} = \frac{n^2}{n^2} \, \frac{3 - \dfrac{5}{n} + \dfrac{2}{n^2}}{1 + \dfrac{7}{n} - \dfrac{4}{n^2}} = \frac{3 - \dfrac{5}{n} + \dfrac{2}{n^2}}{1 + \dfrac{7}{n} - \dfrac{4}{n^2}}.$$

Como
$$\text{lím}\left(3 - \frac{5}{n} + \frac{2}{n^2}\right) = \text{lím } 3 - \text{lím } \frac{5}{n} + \text{lím } \frac{2}{n^2} = 3$$

y
$$\text{lím}\left(1 + \frac{7}{n} - \frac{4}{n^2}\right) = \text{lím } 1 + \text{lím } \frac{7}{n} - \text{lím } \frac{4}{n^2} = 1,$$

$$\text{lím } \frac{3n^2 - 5n + 2}{n^2 + 7n - 4} = \frac{\text{lím}\left(3 - \dfrac{5}{n} + \dfrac{2}{n^2}\right)}{\text{lím}\left(1 + \dfrac{7}{n} - \dfrac{4}{n^2}\right)} = \frac{3}{1} = 3.$$

3.4 Ejemplo. Pruébese que lím $n^a = 0$, si a es un número racional negativo.

SOLUCIÓN. Sea $b = -a$. Entonces $n^a = \left(\dfrac{1}{n}\right)^b$ donde b es un número racional

positivo. La función f definida por $f(x) = x^b$ es continua en 0. Además $\lim \frac{1}{n} = 0$. Luego, según el teorema 3.1,

$$\lim n^a = \lim \left(\frac{1}{n}\right)^b = 0^b = 0\,.$$

Si una sucesión no puede escribirse como una combinación de sucesiones convergentes conocidas, a veces podemos determinar su límite comparando la sucesión con sucesiones cuyo comportamiento se conoce. Este método se basa en el siguiente teorema.

3.5 Teorema. *Si para todos los enteros positivos n, $x_n \leqslant y_n \leqslant z_n$ y si* $\lim x_n = b = \lim z_n$, *entonces $\{y_n\}$ converge y* $\lim y_n = b$.

PRUEBA. Sea un $\varepsilon > 0$ cualquiera. Como $\lim x_n = b$, existe un número N_1 tal que

$$b-\varepsilon < x_n < b+\varepsilon \quad \text{siempre que } n > N_1\,.$$

Como $\lim z_n = b$, existe un número N_2 tal que

$$b-\varepsilon < z_n < b+\varepsilon \quad \text{siempre que } n > N_2\,.$$

Sea $N = \text{máx}\,\{N_1, N_2\}$. Entonces

$$b-\varepsilon < x_n \leqslant y_n \leqslant z_n < b+\varepsilon \quad \text{siempre que } n > N,$$

y, por tanto, $\lim y_n = b$.

3.6 Ejemplo. Pruébese que $\lim \frac{\text{sen}\, n}{n} = 0$.

SOLUCIÓN. Para todo n

$$-\frac{1}{n} \leqslant \frac{\text{sen}\, n}{n} \leqslant \frac{1}{n}\,.$$

Además, $\lim\left(-\frac{1}{n}\right) = 0 = \lim\left(\frac{1}{n}\right)$. Por tanto, por el teorema 3.5, $\lim \frac{\text{sen}\, n}{n} = 0$.

Otro útil resultado respecto a la convergencia de una sucesión, conocido como *prueba de la razón*, se deduce también fácilmente del teorema 3.5.

3.7 Teorema. *Si* $\lim \left|\frac{s_{n+1}}{s_n}\right| < 1$, *entonces* $\lim s_n = 0$.

PRUEBA. Supongamos que $\lim \left|\frac{s_{n+1}}{s_n}\right|$ existe y es menor que 1. Sea r un número tal que

$$\lim \left|\frac{s_{n+1}}{s_n}\right| < r < 1 .$$

Entonces existe un número N tal que

$$\left|\frac{s_{n+1}}{s_n}\right| < r \quad \text{siempre que } n > N.$$

Sea p cualquier entero positivo mayor que N. Entonces, $|s_{p+1}| < r|s_p|$, $|s_{p+2}| < r|s_{p+1}| < r^2|s_p|$ y, en general, para cualquier entero positivo k

$$|s_{p+k}| < r^k|s_p|$$

es decir,

$$-r^k|s_p| < s_{p+k} < r^k|s_p| .$$

Como $r \in \langle 0, 1\rangle$, $\lim\limits_{k\to\infty} r^k = 0$. Por tanto, de acuerdo con el teorema 3.5, $\lim\limits_{k\to\infty} s_{p+k} = 0$, y por tanto, $\lim\limits_{n\to\infty} s_n = 0$.

3.8 Ejemplo. Demuéstrese que $\lim \frac{2^n}{n!} = 0$.

SOLUCIÓN. Para $s_n = \frac{2^n}{n!}$, tenemos

$$\lim \left|\frac{s_{n+1}}{s_n}\right| = \lim \left(\frac{2^{n+1}}{(n+1)!}\Big/\frac{2^n}{n!}\right) = \lim \frac{2}{n+1} = 0 .$$

Luego, según el teorema 3.7, $\lim \frac{2^n}{n!} = 0$.

La siguiente observación nos permitirá obtener el límite de algunas sucesiones de números reales inmediatamente partiendo de los resultados que ya conocemos. Si f es una función real de variable real y $\lim\limits_{x\to\infty} f(x) = b$, entonces $\lim\limits_{n\to\infty} f(n) = b$. Puesto que, si $\lim\limits_{x\to\infty} f(x) = b$ y si $\mathscr{S}(b;\varepsilon)$ es una vecindad cualquiera de b, entonces existe un número N tal que, para todos los números reales x mayores que N, $f(x) \in \mathscr{S}(b;\varepsilon)$. En consecuencia, para cualquier entero positivo n mayor que N, $f(x) \in \mathscr{S}(b;\varepsilon)$.[1]

[1] Es claro que se está suponiendo que $f(n)$ está definida. Es decir que $\mathscr{D}_f$ contiene a todos los enteros positivos. [N. del T.]

Puede que el lector recuerde de sus anteriores estudios que

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e, \quad \lim_{x\to\infty}\frac{\ln x}{x^a} = 0$$

si $a > 0$, y $\lim_{x\to\infty}\dfrac{x^n}{b^x} = 0$ si $b > 1$. De donde tenemos:

3.9 $$\lim\left(1+\frac{1}{n}\right)^n = e$$

3.10 $$\lim\frac{\ln n}{n^a} = 0 \quad \text{cuando } a > 0$$

3.11 $$\lim\frac{n}{b^n} = 0 \qquad \text{cuando } b > 1.$$

3.12 Ejemplo. Pruébese que $\lim \sqrt[n]{n} = 1$.

Solución. Podemos escribir: $\sqrt[n]{n} = n^{1/n} = e^{(1/n)\ln n}$. Como $\lim \dfrac{\ln n}{n} = 0$, $\lim e^{(1/n)\ln n} = e^0 = 1$ (teorema 3.1). Por tanto, $\lim \sqrt[n]{n} = 1$.

Problemas

1. Determínese $\lim s_n$ si s_n es

a) $\dfrac{2n+5}{n-3}$ *b*) $\dfrac{n-7}{n^2+2n}$

c) $\dfrac{n^3+2n^2+1}{3n^5-7n^4+5}$ *d*) $\dfrac{3n^4+n^3-6n}{5n^4+12n^2+7}$

e) $\dfrac{1}{\sqrt{n}}$ *f*) $\dfrac{\sqrt[3]{n}+4}{\sqrt{n}-1}$

g) $\dfrac{2^n}{3+5^n}$ *h*) $\tanh n$.

2. Determínese $\lim s_n$ si s_n es

a) $e^{1/n}$ *b*) $\cos\dfrac{n^2-7}{2n^2+3n}$

c) $\ln(n+1)-\ln n$ *d*) $\tan\dfrac{n}{n+2}$.

3. Úsese el teorema 3.1 para probar que lím $\sqrt[n]{a} = 1$, si $a > 0$.

4. Determínese el límite de $\{s_n\}$ si s_n es

a) $\dfrac{5^n}{n!}$ *b*) $\dfrac{n}{3^n}$

c) $\dfrac{n^2+3}{2^n+n}$ *d*) $\dfrac{n!}{n^n}$

e) $\sqrt{n+1} - \sqrt{n}$.

5. Si a es un número real cualquiera demuéstrese que lím $\dfrac{a^n}{n!} = 0$.

6. Si a es un número real cualquiera y si $x \in \langle -1, 1\rangle$, demuéstrese que

$$\text{lím } \frac{a(a-1)\cdots(a-n)}{n!} x^n = 0.$$

7. Si $0 < b < a$, demuéstrese que lím $\sqrt[n]{a^n+b^n} = a$.

8. Determínese lím $\sqrt[n]{n^2}$.

9. *a*) Demuéstrese que lím $\left(1 + \dfrac{1}{2n}\right)^{2n} = e$.

b) Determínese lím $\left(1 + \dfrac{1}{2n}\right)^{n}$.

10. Determínese lím s_n si s_n es

a) $\sqrt[n]{5}$ *b*) $\dfrac{1}{\sqrt[n]{2}}$

c) $\dfrac{n^3}{4^n}$ *d*) $\dfrac{n^5+5n^3}{2^n-3n^2}$

e) $\dfrac{5 + \ln n}{n^2+n}$ *f*) $\dfrac{2^n+n^4}{3^n-n^7}$.

11. Supongamos que f es una función continua sobre R y sean $f^{(n)}$ la n-ésima iteración de f; es decir, $f^{(1)} = f$ y $f^{(n)} = f \circ f^{(n-1)}$, $n > 1$. Si para algún $a \in$ R, $\lim\limits_{n\to\infty} f^{(n)}(a) = b$, demuéstrese que b es una solución de la ecuación $f(x) = x$.

4. DIVERGENCIA HACIA ∞ O HACIA $-\infty$

La sucesión $\{s_n\}$ donde $s_n = n$ no converge; los términos de esta sucesión no tienden a ningún número, sino que en lugar de ello se hacen indefinidamente grandes. En tal caso decimos que lím $s_n = \infty$.

4.1 Definición. lím $s_n = \infty$ *si para cada número* $K > 0$ *existe un número* N *tal que* $s_n > K$ *siempre que* $n > N$.[1]

Reformulando la definición en términos de vecindades tenemos: lím $s_n = \infty$ si, para cualquier vecindad $\langle K, \infty\rangle$ de ∞, existe una vecindad $\langle N, \infty\rangle$ de ∞ tal que $s_n \in \langle K, \infty\rangle$ siempre que $n \in \langle N, \infty\rangle$.

De un modo análogo definimos lím $s_n = -\infty$.

4.2 Definición. lím $s_n = -\infty$ *si para cada número* $K < 0$ *existe un número* N *tal que* $s_n < K$ *siempre que* $n > N$.[2]

Si lím $s_n = \pm\infty$, entonces decimos que $\{s_n\}$ *diverge a* $\pm\infty$.

4.3 Ejemplo. Demuéstrese que lím $r^n = \infty$ si $r > 1$.

SOLUCIÓN. Tómese $K > 0$. Deseamos demostrar que existe un número N tal que $r^n > K$ siempre que $n > N$. Ahora bien, $r^n > K$ si $n \ln r > \ln N$, es decir, si $n > \dfrac{\ln K}{\ln r}$. (Nótese que $\ln r > 0$ puesto que $r > 1$.) Así pues, si $N = \dfrac{\ln K}{\ln r}$, entonces $r^n > K$ siempre que $n > N$.

Este resultado pudo también haberse obtenido usando siguiente teorema.

4.4 Teorema. *Si para todo* n, $s_n > 0$ *y* lím $s_n = 0$, *entonces* lím $\dfrac{1}{s_n} = \infty$.[3]

PRUEBA. Sea $K > 0$ cualquiera. Deseamos probar que existe un número N tal que $\dfrac{1}{s_n} > K$ siempre que $n > N$. Como lím $s_n = 0$, existe un número N tal

[1] Puede prescindirse de la condición $K > 0$. [N. del T.]

[2] Puede prescindirse de la condición $K < 0$. [N. del T.]

[3] Es frecuente denotar por $+\infty$ (con el signo expreso) el límite que aquí aparece indicado por ∞. En tal caso se dice que lím $s_n = \infty$ (sin ningún signo) cuando (bien sea $\{s_n\}$ una sucesión real o una sucesión compleja) lím $|s_n| = +\infty$. Nótese que lím $s_n = +\infty$ o lím $s_n = -\infty$ implica lím $s_n = \infty$. Con estos convenios son válidas las implicaciones:

i) lím $s_n = 0$, implica lím $\dfrac{1}{s_n} = \infty$;

ii) lím $s_n = \infty$, implica lím $\dfrac{1}{s_n} = 0$. [N. del T.]

que $s_n < \dfrac{1}{K}$ siempre que $n > N$. Usando el hecho de que $s_n > 0$, tenemos, por tanto, que $\dfrac{1}{s_n} > K$ siempre que $n > N$.

4.5 Ejemplo. Demuéstrese que lím $n^a = \infty$ si a es un número racional positivo.

SOLUCIÓN. Sea $b = -a$. Entonces $n^a = \dfrac{1}{n^b}$ donde b es un número racional negativo. Como lím $n^b = 0$ (ejemplo 3.4) y $n^b > 0$, lím $n^a = \infty$ según el teorema 4.4

Combinando los resultados de los ejemplos 3.4 y 4.5 y usan do el hecho de que lím $1 = 1$, tenemos, para cualquier número racional a,

$$\text{lím } n^a = \begin{cases} 0 & \text{si} \quad a < 0 \\ 1 & \text{si} \quad a = 0 \\ \infty & \text{si} \quad a > 0. \end{cases}$$

Podríamos dar bastantes teoremas que enunciasen propiedades de límites infinitos de sucesiones análogas a las propiedades de límites finitos dadas en el corolario 3.2, pág. 458, pero aquí solamente enunciaremos dos más y en la lista de problemas al final de esta sección daremos algunos otros.

4.6 Teorema. *Si* lím $x_n = \infty$ *y* lím $y_n = y > 0$, *entonces* lím $(x_n y_n) = \infty$.

PRUEBA. Tomemos $K > 0$. Como lím $y_n = y > 0$, existe un número N_1 tal que $y_n > \frac{1}{2}y$ siempre que $n > N_1$. Además, como lím $x_n = \infty$, existe un número N_2 tal que $x_n > \dfrac{2K}{y}$ siempre que $n > N_2$. Sea $N = \text{máx}\,\{N_1, N_2\}$. Entonces

$$x_n y_n > \frac{2K}{y} \cdot \frac{y}{2} = K \quad \text{siempre que } n > N$$

y, por tanto, lím $(x_n y_n) = \infty$.

4.7 Teorema. *Si* lím $x_n = \infty$ *y* lím $y_n = y < 0$, *entonces* lím $(x_n y_n) = -\infty$.

PRUEBA. La prueba es análoga a la del teorema 4.6 y se deja para el estudiante.

Estamos ahora en posición de determinar el límite de cualquier sucesión racional.

4.8 Ejemplo. Si $s_n = \dfrac{a_0 + a_1 n + \cdots + a_p n^p}{b_0 + b_1 n + \cdots + b_q n^q}$, donde $a_p \neq 0$ y $b_q \neq 0$, pruébese que

$$\lim s_n = \begin{cases} 0 & \text{si} \quad q > p \\ a_p/b_q & \text{si} \quad q = p \\ \infty & \text{si} \quad q < p \text{ y } a_p b_q > 0 \\ -\infty & \text{si} \quad q < p \text{ y } a_p b_q < 0. \end{cases}$$

SOLUCIÓN

$$s_n = \frac{a_0 + a_1 n + \cdots + a_p n^p}{b_0 + b_1 n + \cdots + b_q n^q} = n^{p-q} \frac{a_0/n^p + a_1/n^{p-1} + \cdots + a_p}{b_0/n^q + b_1/n^{q-1} + \cdots + b_q}.$$

Por el corolario 3.2

$$\lim \frac{a_0/n^p + a_1/n^{p-1} + \cdots + a_p}{b_0/n^q + b_1/n^{q-1} + \cdots + b_q} = \frac{a_p}{b_q}.$$

Entonces, si $q > p$, $\lim n^{p-q} = 0$ y $\lim s_n = 0$ (corolario 3.2). Si $p = q$, $\lim n^{p-q} = 1$ y $\lim s_n = a_p/b_q$ (corolario 3.2). Si $q < p$, $\lim n^{p-q} = \infty$ y $\lim s_n = \infty$ si $a_p/b_q > 0$ y $\lim s_n = -\infty$ si $a_p/b_q < 0$ (teoremas 4.6 y 4.7).

Cuando en el cálculo se introducen las funciones trascendentes ln y exp, es habitual probar que $\lim\limits_{x\to\infty} \ln x = \infty$ y $\lim\limits_{x\to\infty} e^x = \infty$. Admitiendo esto, tenemos,

4.9 $$\lim \ln n = \infty$$

4.10 $$\lim e^n = \infty.$$

En algunos problemas es útil tener un teorema análogo al teorema 3.1, pág. 457, en donde la continuidad de f en $\mathbf{p}$ está reemplazada por una condición sobre el límite de f en $\mathbf{p}$. En el enunciado abajo dado, p y q pueden ser o números reales o $\pm\infty$.

4.11 Teorema. *Supongamos que f es una función de* R *en* R *y que p es un punto de acumulación de $\mathscr{D}^f$. Si* $\lim\limits_p f = q$ *y $\{s_n\}$ es una sucesión de puntos en $\mathscr{D}^f$ tal que, para toda n, $s_n \neq p$ pero* $\lim s_n = p$, *entonces* $\lim f(s_n) = q$.

La prueba de este teorema es análoga a la del teorema 3.1 y la omitimos.

4.12 Ejemplo. Demuéstrese que si a es un número real positivo, entonces $\lim n^a = \infty$.

SOLUCIÓN. Por definición, $n^a = e^{a \ln n}$. Como lím $a \ln n = \infty$ y $\lim_{x\to\infty} e^x = \infty$, tenemos lím $n^a = \infty$ según el teorema 4.11.

Problemas

1. Determínese lím s_n si s_n es

a) 2^n

b) 2^{-n}

c) $\dfrac{3n-2}{5n+12}$

d) $\dfrac{n^3-3n^2+8n-6}{20n^2+32n+15}$

e) $\dfrac{n^5-12n^4}{-2n^2+3n-1}$

f) $\dfrac{12n+3}{n^2-2n+7}$

g) $\dfrac{n\sqrt{n}+4}{n-3}$

h) $\dfrac{n^2+5n}{\sqrt{n}-7}$.

2. Pruébese que lím $\dfrac{b^n}{n^a} = \infty$ cuando $b > 1$.

3. Determínese lím s_n si

a) $s_n = \dfrac{3^n - n^2}{n^5+2}$

b) $s_n = \dfrac{\ln n + 5}{n^2+n}$.

4. Demuéstrese que si lím $x_n = \infty$ y lím $y_n = y$ (un número real), entonces lím $(x_n + y_n) = \infty$.

5. Demuéstrese que si lím $x_n = \infty$ y lím $y_n = \infty$, entonces lím $(x_n + y_n) = \infty$.

6. Si lím $x_n = \infty$ y lím $y_n = -\infty$, entonces, pruébese mediante ejemplos, que todo lo que sigue puede ocurrir:

a) lím $(x_n + y_n) = \infty$

b) lím $(x_n + y_n) = b$ (un número real)

c) lím $(x_n + y_n) = -\infty$

d) lím $(x_n + y_n) \neq b, \pm\infty$.

7. Si $\{s_{n_k}\}$ es una subsucesión de $\{s_n\}$ y $\lim_{n\to\infty} s_n = \infty$, pruébese que $\lim_{k\to\infty} s_{n_k} = \infty$.

8. Pruébese que

a) $\lim_{k\to\infty} \ln 2k = \infty$

b) $\lim_{k\to\infty} \ln (k+1) = \infty$.

9. Si lím $s_n = \infty$, pruébese que lím $\dfrac{1}{s_n} = 0$.

10. Usando el hecho de que $\lim \frac{a^n}{n!} = 0$ para todo número real a, pruébese directamente, partiendo de la definición, que $\lim \sqrt[n]{n!} = \infty$.

11. Si $a_n \geqslant b_n$ y $\lim b_n = \infty$, pruébese que $\lim a_n = \infty$.

12. Pruébese el teorema 4.11.

5. SUCESIONES MONÓTONAS

La definición de una sucesión monótona está comprendida en la definición de una función monótona. Sin embargo, enunciamos aquí las definiciones pertinentes para este caso particular que son las sucesiones.

5.1 Definición. *Una sucesión* $\{s_n\}$ *es* ***no decreciente (no creciente)*** *si* $s_n \leqslant s_{n+1}$ $(s_n \geqslant s_{n+1})$ *para todo* n.

Si la desigualdad se verifica siempre en la definición anterior, es decir, si $s_n < s_{n+1}$ $(s_n > s_{n+1})$ para todo n, entonces decimos que $\{s_n\}$ es *creciente (decreciente)*. Claramente, una sucesión creciente es no decreciente y una decreciente es no creciente.[1]

5.2 Definición. *Una sucesión es* ***monótona*** *si es no decreciente o no creciente.*

Es fácil imaginar el comportamiento de una sucesión monótona. Por ejemplo, supongamos que $\{s_n\}$ es una sucesión creciente que está acotada superiormente por el número b. Entonces, los términos de la sucesión se

FIGURA 2

mueven a la derecha sobre la recta de los números (figura 2), pero nunca llegan a estar a la derecha de b. En realidad, los términos nunca llegan a estar a la derecha del supremo de $\{s_n\}$, pero sí llegan a estar tan cerca como se desee de este número. Parece, por ello, que esta sucesión habría de convergir a su supremo. Esto es lo que se prueba en el siguiente teorema.

5.3 Teorema. *Cualquier sucesión monótona acotada* $\{s_n\}$ *converge. Si* $\{s_n\}$ *es no decreciente (no creciente) entonces* $\lim s_n = \sup \{s_n\}$ $(\inf \{s_n\})$.

[1] Algunos autores prefieren llamar "crecientes" a las sucesiones que aquí se llaman "no decrecientes", y "decrecientes" a las que aquí se llaman "no crecientes". Cuando tal es el caso, a las que aquí se llaman "crecientes" se les llama "estrictamente crecientes", y a las "decrecientes", "estrictamente decrecientes". [N. del T.]

PRUEBA. Supongamos que $\{s_n\}$ es no decreciente y $s_n \leqslant b$ para todo n. Como los términos de la sucesión constituyen un conjunto no nulo de números reales que está superiormente acotado, este conjunto tiene un supremo, llamémosle c. Probamos ahora que lím $s_n = c$. Tomemos $\varepsilon > 0$. El número $c-\varepsilon$ no puede ser una cota superior de $\{s_n\}$. Por tanto, para algún entero N, $s_N > c-\varepsilon$. Como $\{s_n\}$ es no decreciente,

$$s_n > c-\varepsilon \quad \text{para toda } n > N.$$

Por otra parte, para todo n, $s_n \leqslant c$. Por tanto,

$$s_n \in \langle c-\varepsilon, c+\varepsilon \rangle \quad \text{para toda } n > N$$

y, por tanto, lím $s_n = c$.

Si $\{s_n\}$ es no creciente y acotada, entonces $\{-s_n\}$ es no decreciente y acotada y, por tanto, converge. De esta manera,

$$\text{lím } s_n = \text{lím } [-(-s_n)] = -\text{lím } (-s_n) = -\sup \{-s_n\} = \inf \{s_n\}.$$

Y esto completa la prueba.

Si $\{s_n\}$ es no decreciente, pero no acotada, es decir, no acotada superiormente, entonces $\{s_n\}$ diverge a ∞. Si $\{s_n\}$ es no creciente, pero no acotada, entonces $\{s_n\}$ diverge a $-\infty$. (La fácil prueba de estas afirmaciones se deja como ejercicio para el estudiante.)

Tenemos, pues, el siguiente corolario del teorema 5.3.

5.4 Corolario. *Una sucesión no decreciente (no creciente) o converge o diverge a* $\infty (-\infty)$ *según que sea acotada o no.*

Nota. Como el límite de una sucesión depende solamente del comportamiento de los términos de la sucesión a partir de algún término en adelante, el anterior teorema se verifica también para sucesiones que son monótonas a partir de un término cualquiera.

5.5 Ejemplo. Demuéstrese que lím ln $n = \infty$.

SOLUCIÓN. Como $D_x \ln x = \dfrac{1}{x} > 0$ para $x > 0$, la función logarítmica es una función creciente y, por tanto, $\{\ln n\}$ es una sucesión creciente. Por tanto, según el corolario 5.4, lím ln $n = p$, donde p es un número real o ∞. La sucesión $\{\ln 2n\}$ es una subsucesión de la ln x y, por tanto, lím ln $2n = p$. Como

$$\ln 2n = \ln 2 + \ln n,$$

si p fuera un número real tendríamos

$$p = \ln 2 + p$$

lo que es imposible. Por tanto, $p = \infty$.

5.6 Ejemplo. Pruébese que la sucesión

$$\sqrt{2},\ \sqrt{2\sqrt{2}},\ \sqrt{2\sqrt{2\sqrt{2}}},\ \ldots$$

converge a 2.

SOLUCIÓN. Damos primero una descripción más explícita de la sucesión:

$$s_1 = \sqrt{2}$$
$$s_n = \sqrt{2s_{n-1}}, \qquad n > 1.$$

Probaremos ahora que $\{s_n\}$ es una sucesión no decreciente acotada superiormente por 2. La prueba es por inducción matemática. Sea $\mathscr{S}$ el conjunto de los enteros positivos n tales que $s_n \leqslant 2$ y $s_n \leqslant s_{n+1}$.

1) $1 \in \mathscr{S}$ ya que $s_1 = \sqrt{2} < 2$ y $s_1 = \sqrt{2} < \sqrt{2\sqrt{2}} = s_2$.

2) Supongamos que $m \in \mathscr{S}$, es decir, que $s_m \leqslant 2$ y $s_m \leqslant s_{m+1}$. Entonces

$$s_{m+1} = \sqrt{2s_m} \leqslant \sqrt{4} = 2$$

y

$$s_{m+1} = \sqrt{s_{m+1}^2} \leqslant \sqrt{2s_{m+1}} = s_{m+2}.$$

Por tanto, $m \in \mathscr{S}$ implica $m+1 \in \mathscr{S}$.

$\mathscr{S}$ es, pues, el conjunto de todos los enteros positivos de acuerdo con el principio de inducción, y hemos mostrado que $\{s_n\}$ es una sucesión no decreciente acotada. Luego $\{s_n\}$ converge. Sea lím $s_n = c$. Como $s_{n+1} = \sqrt{2s_n}$ y lím $s_{n+1} = c$, tenemos $c = \sqrt{2c}$. Por tanto, $c^2 = 2c$ y $c(c-2) = 0$. c es, pues, o 0 o 2. Pero como $\{s_n\}$ es una sucesión de términos positivos, 0 no puede ser el supremo de $\{s_n\}$ y, por tanto, $c = 2$.

Problemas

1. Pruébese que lím $e^n = \infty$.

2. Pruébese que $\dfrac{\ln x}{x^a}$ (donde $a > 0$) es decreciente sobre algún intervalo $\langle x_0, \infty\rangle$ y pruébese luego que lím $\dfrac{\ln n}{n^a} = 0$.

3. Determínese el límite de la siguiente sucesión: $\sqrt{2}$, $\sqrt{2+\sqrt{2}}$, $\sqrt{2+\sqrt{2+\sqrt{2}}}, \ldots$

4. Si $\{s_n\}$ es una sucesión no decreciente que no es acotada, pruébese que lím $s_n = \infty$.

5. Si $f(x) = \frac{1}{2}\left(x + \frac{2}{x}\right)$ y $s_{n+1} = f(s_n)$, determínese lím s_n cuando

a) $s_1 = 2$ *b*) $s_1 = 1$ *c*) $s_1 = -1$ *d*) $s_1 = a$ $(a \neq 0)$.

6. Si f es una función real de variable real que es no decreciente sobre el intervalo $[1, \infty\rangle$, pruébese que $\lim_{x \to \infty} f(x) = b$ (un número real) o $\lim_{x \to \infty} f(x) = \infty$.

6. PUNTOS LÍMITES DE UNA SUCESIÓN

La sucesión $\{(-1)^n\}$ ni converge ni diverge a $\pm\infty$. Hay una infinidad de términos de esta sucesión "próximos" a 1 y también una infinidad "próximos" a -1. Decimos que 1 y -1 son puntos límites de la sucesión.[1]

6.1 Definición. *Un punto p (número real o $\pm\infty$) es un* ***punto límite*** *(o punto de acumulación) de una sucesión $\{s_n\}$ de números reales si toda vecindad de p contiene infinitos términos de la sucesión.*

Es decir, si el punto límite p es finito (un número real), entonces para cualquier $\varepsilon > 0$ y para cualquier entero positivo m existe un entero $n > m$ tal que $s_n \in \mathscr{S}(p; \varepsilon) = \langle p-\varepsilon, p+\varepsilon\rangle$. Si el punto límite p es $\infty(-\infty)$ entonces, para cualquier número a y cualquier entero positivo m, existe un entero $n > m$ tal que $s_n \in \langle a, \infty\rangle$ $(\langle -\infty, a\rangle)$.

6.2 Ejemplo. Pruébese que 1 y -1 son los únicos puntos límites de la sucesión $\{(-1)^n\}$.

SOLUCIÓN. Los puntos 1 y -1 son puntos límites de la sucesión: cualquier vecindad de 1 contiene infinitos términos de la sucesión —todos los términos pares. Todos los términos impares de la sucesión se encuentran en cualquier vecindad de -1. Ningún otro punto es un punto límite ya que podemos encontrar una vecindad del punto que no contiene ni 1 ni -1 y, por tanto, no contienen ningún término de la sucesión.

Usaremos el término "punto límite" con preferencia al de "punto de acumulación", aunque este último es muy sugestivo: un número infinito de términos de la sucesión se acumulan alrededor del punto. Sin embargo, debe advertirse que un punto de acumulación de una sucesión no es necesariamente un punto de acumulación del rango de la sucesión. El rango de la sucesión $\{(-1)^n\}$ es el conjunto $\{-1, 1\}$ que no tiene puntos de

[1] En otras terminologías también en uso, a los que aquí se llaman puntos límites se les llama "límites de oscilación". [N. del T.]

acumulación. Por otra parte, si un punto es un punto de acumulación del rango de una sucesión, entonces es un punto límite o de acumulación de la sucesión.

Debe tenerse cuidado en no confundir los términos "punto límite de una sucesión" y "límite de una sucesión". Un punto es el límite de una sucesión si cualquier vecindad del punto contiene a todos los términos de la sucesión, salvo un número finito. Por otra parte, un punto es un punto límite de una sucesión si cualquier vecindad del punto contiene infinitos términos de la sucesión. Claramente, el límite de una sucesión es un punto límite de una sucesión. Sin embargo, el recíproco no es cierto necesariamente, ya que una vecindad puede contener infinitos términos de una sucesión y, al mismo tiempo, puede haber infinitos términos de la sucesión que no estén en la vecindad. Por ejemplo, la vecindad $\langle 0, 2\rangle$ de 1 contiene todos los términos pares, pero ninguno de los impares de la sucesión $\{(-1)^n\}$.

El número de puntos límites de una sucesión caracteriza su comportamiento respecto a la convergencia. Toda sucesión tiene al menos un punto límite (que puede ser $+\infty$ o $-\infty$). Si una sucesión tiene solamente un punto límite y es finito (es decir, un número real), entonces la sucesión converge a ese punto. Si una sucesión tiene ∞ o $-\infty$ como su solo punto límite, entonces la sucesión diverge a ∞ o $-\infty$, respectivamente. Si una sucesión tiene más de un punto límite, entonces ni converge ni diverge a $\pm\infty$; decimos que la sucesión *oscila* (o que es *oscilante*).

En seguida probaremos estas afirmaciones.

6.3 Teorema. *Cualquier sucesión $\{s_n\}$ de números reales tiene al menos un punto límite.*

PRUEBA. Sea $\mathscr{S}$ el conjunto de números reales x tales que $x \leqslant s_n$ para infinitas n. Si el conjunto $\mathscr{S}$ es vacío, entonces $-\infty$ es un punto límite de $\{s_n\}$ ya que todos, salvo un número finito de términos de $\{s_n\}$, se encuentran en cualquier vecindad $\langle -\infty, x\rangle$ de $-\infty$. Si $\mathscr{S}$ no está superiormente acotado, entonces infinitos términos de $\{s_n\}$ se encuentran en cualquier vecindad $\langle x, \infty\rangle$ de ∞ y, por tanto, ∞ es un punto límite de $\{s_n\}$. Si $\mathscr{S}$ es no vacío y superiormente acotado, entonces $\mathscr{S}$ tiene un supremo, llamémosle c. Probemos que c es un punto límite de $\{s_n\}$. En cualquier vecindad $\langle c-\varepsilon, c+\varepsilon\rangle$ de c hay un punto $x \in \mathscr{S}$. Como $x \in \mathscr{S}$, $x \leqslant s_n$ para infinitos n. Por otra parte, $c+\varepsilon \notin \mathscr{S}$ y, por tanto, $s_n \geqslant c+\varepsilon$ para solamente un número finito de valores de n. Así pues, infinitos términos de $\{s_n\}$ se encuentran en $\langle c-\varepsilon, c+\varepsilon\rangle$. Esto prueba que c es un punto límite de $\{s_n\}$ y completa la prueba.

Como una sucesión acotada no puede tener $\pm\infty$ como punto límite, podemos enunciar: *una sucesión acotada tiene, al menos, un punto límite finito.*

6.4 Teorema. lím $s_n = p$ *(donde p puede ser tanto un real como* $\pm\infty$*) si y sólo si* $\{s_n\}$ *tiene p como único punto límite.*

PRUEBA. Supongamos que lím $s_n = p$. Entonces p es un punto límite de $\{s_n\}$. Sea q un punto cualquiera distinto de p. Tómese una vecindad cualquiera $\mathcal{N}_p$ de p y una vecindad cualquiera $\mathcal{N}_q$ de q tales que $\mathcal{N}_p \cap \mathcal{N}_q = \varnothing$. Como lím $s_n = p$, todos, salvo un número finito de términos de $\{s_n\}$, se encuentran en $\mathcal{N}_p$. Por tanto, la vecindad $\mathcal{N}_q$ de q no puede contener infinitos términos de $\{s_n\}$. Esto prueba, que q no puede, ser un punto límite de $\{s_n\}$ y, por tanto, si lím $s_n = p$, $\{s_n\}$ puede tener solamente a p como punto límite.

Supongamos que $\{s_n\}$ tiene a p como su único punto límite. Deseamos demostrar que lím $s_n = p$. Tomemos una vecindad cualquiera $\mathcal{N}_p$ de p. Supongamos que hay un número infinito de términos de $\{s_n\}$ que no están en $\mathcal{N}_p$. Entonces, estos términos constituyen una subsucesión de $\{s_n\}$ y esta subsucesión tiene un punto límite según el teorema 6.3. Este punto límite es un punto límite de la sucesión $\{s_n\}$ distinto de p. Esto contradice el hecho de que p es el único punto límite de $\{s_n\}$. Por tanto, todos los términos de $\{s_n\}$, salvo cuando más un número finito de ellos, deben encontrarse en cualquier vecindad de p y, por tanto, lím $s_n = p$. Lo que completa la prueba.

6.5 Ejemplo. Pruébese que la sucesión $\{r^n\}$, donde $r < -1$, oscila.

SOLUCIÓN. Probaremos que $\{r^n\}$ tiene los dos puntos límites ∞ y $-\infty$. Tómese $K > 0$. Como $|r| > 1$, lím $|r|^n = \infty$ (ejemplo 4.3, pág. 463). Por tanto, existe un número N tal que $|r|^n \in \langle K, \infty\rangle$ siempre que $n > N$. Así pues, si n es par y $n > N$, $r^n \in \langle K, \infty\rangle$ y, si n es impar y $n > N$, $r^n \in \langle -\infty, -K\rangle$. Esto prueba que cualquier vecindad de ∞ y cualquier vecindad de $-\infty$ contienen infinitos términos de la sucesión y, por tanto, que ∞ y $-\infty$ son puntos límites de $\{r^n\}$. Así pues, $\{r^n\}$, cuando $r < -1$, no converge ni diverge a $\pm\infty$; oscila.

Combinando el resultado del ejemplo 6.5 con los resultados previamente derivados, podemos enunciar; $\{r^n\}$ converge a 0 si $|r| < 1$, converge a 1 si $r = 1$, diverge a ∞ si $r > 1$, y oscila si $r \leqslant -1$.

Ahora estableceremos una relación entre puntos límites y subsucesiones de una sucesión.

6.6 Teorema. *Un punto p (un número real o* $\pm\infty$*) es un punto límite de la sucesión* $\{s_n\}$ *si y sólo si hay una subsucesión de la* $\{s_n\}$ *que converge a p si p es un número real o diverge a p si* $p = \pm\infty$.

PRUEBA. Sea $\{s_{n_k}\}$ una subsucesión de la sucesión $\{s_n\}$ y sea lím $s_{n_k} = p$. Sea $\mathcal{N}$ una vecindad cualquiera de p. Todos los términos de $\{s_{n_k}\}$, salvo un

número finito, se encuentran en $\mathcal{N}$. Luego, de aquí que en $\mathcal{N}$ se encuentran infinitos términos de $\{s_n\}$. Luego p es un punto límite de $\{s_n\}$.

Supongamos que el número real p es un punto límite de $\{s_n\}$. Entonces toda vecindad de p contiene infinitos términos de $\{s_n\}$. Podemos escoger una subsucesión $\{s_{n_k}\}$ que converja a p como sigue. Sea s_{n_1} un término cualquiera de $\{s_n\}$ que pertenezca a $\mathcal{S}(p; 1)$. Sea s_{n_2} un término cualquiera posterior a s_{n_1} en $\{s_n\}$ que pertenezca a $\mathcal{S}(p; \frac{1}{2})$. Sea s_{n_3} un término cualquiera posterior a s_{n_2} en $\{s_n\}$ que pertenezca a $\mathcal{S}(p; \frac{1}{3})$. Si continuamos escogiendo términos de esta forma, obtenemos una subsucesión $\{s_{n_k}\}$ que converge a p.

Si ∞ es un punto límite de $\{s_n\}$, entonces escogemos una subsucesión $\{s_{n_k}\}$ que diverge a ∞ como sigue. Tomamos como s_{n_1} cualquier término de $\{s_n\}$ que se encuentre en $\langle 1, \infty\rangle$. s_{n_2} es cualquier término posterior a s_{n_1} en $\{s_n\}$ que se encuentre en $\langle 2, \infty\rangle$. Si continuamos nuestra elección de términos de este modo obtenemos una subsucesión $\{s_{n_k}\}$ que diverge a ∞. Si es $-\infty$ el que es un punto límite de $\{s_n\}$, podemos escoger una subsucesión $\{s_{n_k}\}$ que diverja a $-\infty$ de modo análogo. Esto completa la prueba.

Sea $\mathcal{S}$ el conjunto de todos los puntos límites de una sucesión $\{s_n\}$. Probaremos que $\mathcal{S}$ tiene máximo y mínimo. (Cuando introdujimos los puntos ideales ∞ y $-\infty$ establecimos que $-\infty < x < \infty$ para todo número real x.) Sea p el supremo de $\mathcal{S}$; si $\mathcal{S}$ no está superiormente acotado, entonces $p = \infty$. Probaremos ahora que p es un punto límite de $\{s_n\}$ y, por tanto, que p es el máximo de $\mathcal{S}$. Sea $\mathcal{N}$ una vecindad cualquiera de p. Como p es el supremo de $\mathcal{S}$, hay un punto límite b de $\{s_n\}$ en $\mathcal{N}$. Si $\mathcal{M}$ es una vecindad de b tal que $\mathcal{M} \subset \mathcal{N}$, entonces hay infinitos términos de $\{s_n\}$ en $\mathcal{M}$, luego en $\mathcal{N}$. Así pues, toda vecindad de p contiene infinitos términos de $\{s_n\}$ y p es un punto límite de $\{s_n\}$. Prueba esto que p es el máximo punto límite de $\{s_n\}$. Podríamos probar en forma análoga que $\{s_n\}$ tiene un punto límite mínimo.

6.7 Definición. *El **límite superior** de una sucesión $\{s_n\}$, denotado por* $\overline{\lim}\, s_n$, *es el máximo punto límite de $\{s_n\}$.*

6.8 Definición. *El **límite inferior** de una sucesión $\{s_n\}$, denotado por* $\underline{\lim}\, s_n$, *es el punto límite mínimo de $\{s_n\}$.*

Reformularemos ahora el teorema 6.4 en términos de límites superior e inferior.

6.9 Teorema. $\lim s_n = p$ *(donde p puede ser $\pm\infty$) si y sólo si* $\underline{\lim}\, s_n = p = \overline{\lim}\, s_n$.

6.10 Ejemplo. ¿Converge la sucesión $\left\{\operatorname{sen} \dfrac{n\pi}{4}\right\}$?

SOLUCIÓN. Nótese, en primer término, que $-1 \leqslant \operatorname{sen} \frac{n\pi}{4} \leqslant 1$ para todo n. Si n es de la forma $8k+2$, donde k es un entero, entonces $\operatorname{sen} \frac{n\pi}{4} = 1$. Vemos, pues, que 1 es un punto límite de la sucesión, en realidad, $\overline{\lim} \operatorname{sen} \frac{n\pi}{4} = 1$. Si n es de la forma $8k+6$, entonces $\operatorname{sen} \frac{n\pi}{4} = -1$. Por tanto, $\underline{\lim} \operatorname{sen} \frac{n\pi}{4} = -1$. Como los límites superior e inferior de la sucesión son diferentes, la sucesión oscila.

6.11 Ejemplo. Si $\lim x_n = x > 0$ y $\overline{\lim}\, y_n = y$ (donde $x, y \in \mathrm{R}$), pruébese que $\overline{\lim}\, x_n y_n = xy$.

SOLUCIÓN. Como $\overline{\lim}\, y_n = y$, y es un punto límite de $\{y_n\}$ y, por tanto, existe una subsucesión $\{y_{n_k}\}$ tal que $\lim y_{n_k} = y$. Además, $\lim x_{n_k} = x$ ya que $\lim x_n = x$. Por tanto, $\lim \{x_{n_k} y_{n_k}\} = xy$. Prueba esto que xy es un punto límite de $\{x_n y_n\}$. Supongamos que $\{x_n y_n\}$ tiene un punto límite mayor que xy, llamémosle p. Entonces existe una subsucesión $\{x_{n_j} y_{n_j}\}$ tal que $\lim x_{n_j} y_{n_j} = p$. Si p es finito, entonces

$$\lim y_{n_j} = \lim \left(\frac{1}{x_{n_j}} x_{n_j} y_{n_j} \right) = \frac{p}{x}.$$

Si $p = \infty$, entonces

$$\lim y_{n_j} = \lim \left(\frac{1}{x_{n_j}} x_{n_j} y_{n_j} \right) = \infty .$$

En cualquier caso $\{y_n\}$ tendría un punto límite mayor que y. Por tanto, xy es el máximo punto límite de $\{x_n y_n\}$.

Podemos caracterizar el límite superior de una sucesión como sigue.

6.12 Teorema. $\overline{\lim}\, s_n = p$ *si y sólo si para cualquier vecindad* $\mathcal{N}_p$ *de* p, *se tiene: el número de términos de* $\{s_n\}$ *en* $\mathcal{N}_p$ *es infinito y el número de términos de* $\{s_n\}$ *a la derecha de* $\mathcal{N}_p$ *es finito (posiblemente cero).*

PRUEBA. Supongamos que hay infinitos términos de $\{s_n\}$ en cualquier vecindad $\mathcal{N}_p$ de p y, también cualquier vecindad de p, $\mathcal{N}_p$, tan solo un número finito de términos a su derecha. Tomemos un punto cualquiera $q > p$. Consideremos vecindades $\mathcal{N}_p$ de p y $\mathcal{N}_q$ de q tales que $\mathcal{N}_p \cap \mathcal{N}_q = \varnothing$. Como $\mathcal{N}_q$ está a la derecha de $\mathcal{N}_p$ no puede contener infinitos términos

de $\{s_n\}$ y, por tanto, q no puede ser un punto límite de $\{s_n\}$. Luego p es el máximo punto límite de $\{s_n\}$.

Supongamos que $\overline{\lim}\, s_n = p$. Entonces cualquier vecindad $\mathcal{N}_p$ de p contiene infinitos términos de $\{s_n\}$. Si hubiera infinitos términos de $\{s_n\}$ a la derecha de $\mathcal{N}_p$, entonces estos términos constituirían una subsucesión que tendría un punto límite según el teorema 6.3. Este punto límite sería un punto límite de $\{s_n\}$ mayor que p. Así pues, sólo puede haber un número finito de términos de $\{s_n\}$ a la derecha de $\mathcal{N}_p$. Y esto completa la prueba.

Podemos caracterizar el límite inferior de una sucesión de un modo análogo —simplemente cambiando la palabra "derecha" en el teorema 6.12 por la palabra "izquierda".

Problemas

1. Determínense todos los puntos límites de las siguientes sucesiones $\{s_n\}$.

a) $s_n = (-1)^n\left(2 + \dfrac{1}{n}\right)$ *b*) $s_n = \operatorname{sen}\dfrac{n\pi}{3}$

c) $s_n = (-1)^n n$ *d*) $s_n = n+(-1)^n n$.

2. Si p es un punto de acumulación del rango de una sucesión $\{s_n\}$, pruébese que p es un punto límite de $\{s_n\}$.

3. Si p es un punto límite de una subsucesión $\{s_{n_k}\}$ de $\{s_n\}$, pruébese que p es un punto límite de $\{s_n\}$.

4. Pruébese que $\underline{\lim}\, s_n = -\overline{\lim}\,(-s_n)$.

5. Determínese $\underline{\lim}\, s_n$ y $\overline{\lim}\, s_n$ si s_n es

a) $n \operatorname{sen}\dfrac{n\pi}{3}$ *b*) $\dfrac{1}{n}\operatorname{sen}\dfrac{n\pi}{3}$

c) $\left(1 + \dfrac{1}{n}\right)\operatorname{sen}\dfrac{n\pi}{3}$ *d*) $\dfrac{n+n^2 \operatorname{sen}\dfrac{n\pi}{3}}{n^2+4}$

e) $\dfrac{n}{4} - \left[\dfrac{n}{4}\right]$ donde [ ] es la función "máximo entero".

6. Pruébese que:

a) $\underline{\lim}\, x_n + \underline{\lim}\, y_n \leqslant \underline{\lim}\,(x_n+y_n)$

b) $\overline{\lim}\,(x_n+y_n) \leqslant \overline{\lim}\, x_n + \overline{\lim}\, y_n$.

7. Proporciónese un ejemplo en el que se verifique la desigualdad del problema 6a.

8. Si $\lim x_n = x > 0$ y $\overline{\lim}\ y_n = \infty$, pruébese que $\overline{\lim}\ x_n y_n = \infty$.

9. Si $x_n \leqslant y_n$ para todo n, pruébese que $\underline{\lim}\ x_n \leqslant \underline{\lim}\ y_n$ y $\overline{\lim}\ x_n \leqslant \overline{\lim}\ y_n$.

10. Pruébese que:

a) $\inf \{s_n\} \leqslant \underline{\lim}\ s_n$
b) $\overline{\lim}\ s_n \leqslant \sup \{s_n\}$.

11. Si $s_n = \dfrac{b+a}{2} + (-1)^n \dfrac{b-a}{2}$ donde $a < b$, determínese $\underline{\lim}\ s_n$ y $\overline{\lim}\ s_n$.

12. Proporciónese una sucesión $\{s_n\}$ con los siguientes límites inferior y superior:

a) $\underline{\lim}\ s_n = 2$ y $\overline{\lim}\ s_n = 8$
b) $\underline{\lim}\ s_n = -5$ y $\overline{\lim}\ s_n = 21$.

7. ALGUNOS TEOREMAS SOBRE FUNCIONES CONTINUAS DE UN VECTOR

Si una función real de variable real es continua sobre un intervalo cerrado, entonces la función es uniformemente continua y acotada sobre el intervalo cerrado. Resultados similares se verifican para funciones continuas de R^m en R^n si se reemplaza el intervalo cerrado por un conjunto cerrado y acotado en R^m. Para obtener estos resultados usaremos aquí sucesiones de puntos de R^m.

Dada una sucesión de puntos en R^m definimos un punto límite de la sucesión de modo análogo al empleado para definir un punto límite para una sucesión de números reales: un punto $\mathbf{c}$ es un *punto límite* de una sucesión $\{\mathbf{s}_n\}$ si toda vecindad de $\mathbf{c}$ contiene infinitos términos de la sucesión. Para sucesiones de puntos tenemos resultados análogos a los enunciados en los teoremas 6.6 y 6.3 para sucesiones de números reales.

7.1 Teorema. *Un punto* $\mathbf{c}$ *es un punto límite de la sucesión* $\{\mathbf{s}_n\}$ *si y sólo si hay una subsucesión de* $\{\mathbf{s}_n\}$ *que converge a* $\mathbf{c}$.

La prueba de este teorema es similar a la del teorema 6.6, pág. 472, y por ello no la daremos.

7.2 Teorema. *Toda sucesión acotada* $\{\mathbf{s}_n\}$ *de puntos en* R^m *tiene un punto límite.*

PRUEBA. Probaremos este teorema para el caso en que $\{\mathbf{s}_n\}$ es una sucesión de puntos en R^2. El caso general podría probarse en forma análoga usando

inducción. Sea $\mathbf{s}_n = (x_n, y_n)$. La sucesión $\{x_n\}$ es una sucesión acotada de números reales y, por tanto, tiene un punto límite c. Hay entonces, de acuerdo con el teorema 6.6, una subsucesión $\{x_{n_k}\}$ de $\{x_n\}$ tal que $\lim x_{n_k} = c$. La sucesión $\{y_{n_k}\}$ es una sucesión acotada de números reales y tiene, por tanto, un punto límite d. Hay entonces una subsucesión $\{y_{n_{k_i}}\}$ de $\{y_{n_k}\}$ tal que $\lim y_{n_{k_i}} = d$. Por tanto

$$\lim \mathbf{s}_{n_{k_i}} = \lim (x_{n_{k_i}}, y_{n_{k_i}}) = (c, d) = \mathbf{c}.$$

Luego, según el teorema 7.1, $\mathbf{c}$ es un punto límite de $\{\mathbf{s}_n\}$. Y esto completa la prueba.

Probaremos ahora que si el rango de una sucesión $\{\mathbf{s}_n\}$ se encuentra en un conjunto acotado y cerrado $\mathscr{E}$, entonces $\{\mathbf{s}_n\}$ tiene un punto límite en $\mathscr{E}$. Nuestra prueba consistirá en demostrar que cualquier punto límite de la sucesión $\{\mathbf{s}_n\}$ es o un punto del rango de $\{\mathbf{s}_n\}$ o un punto de acumulación del rango de $\{\mathbf{s}_n\}$ y usar luego el hecho de que un conjunto cerrado contiene a todos sus puntos de acumulación.

7.3 Lema. *Un punto límite* $\mathbf{c}$ *de una sucesión* $\{\mathbf{s}_n\}$ *es o un elemento del conjunto* $\{\mathbf{s}_n\}$ *o un punto de acumulación de este conjunto.*

PRUEBA. Supongamos que $\mathbf{c} \notin \mathbf{s}_n$. Tomemos una vecindad $\mathscr{S}(\mathbf{c}; \varepsilon)$ de $\mathbf{c}$. Como $\mathbf{c}$ es un punto límite de la sucesión $\{\mathbf{s}_n\}$, $\mathscr{S}(\mathbf{c}; \varepsilon)$ contiene infinitos términos de la sucesión. Luego $\mathscr{S}(\mathbf{c}; \varepsilon)$ contiene un punto del conjunto $\{\mathbf{s}_n\}$ y ha de ser distinto de $\mathbf{c}$. Lo que prueba que si $\mathbf{c}$ no es un elemento del conjunto $\{\mathbf{s}_n\}$, entonces $\mathbf{c}$ es un punto de acumulación de ese conjunto.

7.4 Teorema. *Si* $\{\mathbf{s}_n\}$ *es una sucesión de puntos que se encuentran en un conjunto cerrado y acotado* $\mathscr{E}$, *entonces* $\{\mathbf{s}_n\}$ *tiene un punto límite en* $\mathscr{E}$.

PRUEBA. Por el teorema 7.2 sabemos que $\{\mathbf{s}_n\}$ tiene un punto límite $\mathbf{c}$. El punto $\mathbf{c}$ es o un elemento o un punto de acumulación del conjunto $\{\mathbf{s}_n\}$ y, por tanto, del conjunto $\mathscr{E}$ (lema 7.3). Luego $\mathbf{c} \in \bar{\mathscr{E}}$, la cerradura de $\mathscr{E}$, y, como $\mathscr{E}$ es cerrado, $\mathbf{c} \in \mathscr{E}$.

Podemos ahora probar que una función de R^m a R^n que es continua sobre un conjunto cerrado y acotado $\mathscr{E}$ es uniformemente continua sobre $\mathscr{E}$. Primero definimos la continuidad uniforme para una función de un vector.

7.5 Definición. *La función* $\mathbf{f}$ *de* R^m *en* R^n *es* ***uniformemente continua*** *sobre el conjunto* $\mathscr{E}$ *si* $\mathscr{E}$ *está contenido en el dominio de* $\mathbf{f}$ *y si para cualquier* $\varepsilon > 0$ *existe una* $\delta > 0$ *mayor que* 0 *tal que si* $\mathbf{x}$ *y* $\mathbf{y}$ *pertenecen a* $\mathscr{E}$ *y* $|\mathbf{x}-\mathbf{y}| < \delta$, *entonces*

$$|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{y})| < \varepsilon.$$

7.6 Teorema. *Si* **f** *es continua sobre un conjunto cerrado y acotado* $\mathscr{E}$, *entonces* **f** *es uniformemente continua sobre* $\mathscr{E}$.

PRUEBA. Supongamos que **f** no es uniformemente continua sobre $\mathscr{E}$. Existe entonces un número $\varepsilon > 0$ tal que para todos los enteros positivos n existen puntos $\mathbf{x}_n$ y $\mathbf{y}_n$ en $\mathscr{E}$ tales que $|\mathbf{x}_n - \mathbf{y}_n| < 1/n$ y $|\mathbf{f}(\mathbf{x}_n) - \mathbf{f}(\mathbf{y}_n)| > \varepsilon$. La sucesión $\{\mathbf{x}_n\}$ tiene un punto límite **c** en $\mathscr{E}$ y, por tanto, una subsucesión $\{\mathbf{x}_{n_k}\}$ tal que

$$\text{lím } \mathbf{x}_{n_k} = \mathbf{c}.$$

Como $|\mathbf{x}_{n_k} - \mathbf{y}_{n_k}| < 1/n_k$ y $\lim\limits_{k\to\infty} \dfrac{1}{n_k} = 0$,

$$\text{lím } \mathbf{y}_{n_k} = \mathbf{c}.$$

Como **f** es continua sobre $\mathscr{E}$, usando el teorema 3.1, pág. 457, tenemos

$$\text{lím } \mathbf{f}(\mathbf{x}_{n_k}) = \mathbf{f}(\mathbf{c})$$

y

$$\text{lím } \mathbf{f}(\mathbf{y}_{n_k}) = \mathbf{f}(\mathbf{c}).$$

Pero, para todo k, $|\mathbf{f}(\mathbf{x}_{n_k}) - \mathbf{f}(\mathbf{y}_{n_k})| \geqslant \varepsilon$. Esta contradicción prueba que la suposición de que **f** no es uniformemente continua sobre $\mathscr{E}$ no puede verificarse. Lo que completa la prueba.

7.7 Teorema. *Si* **f** *es continua sobre un conjunto cerrado y acotado* $\mathscr{E}$ *entonces* **f** *es acotada sobre* $\mathscr{E}$.

PRUEBA. Supongamos que **f** no fuera acotada sobre $\mathscr{E}$. Entonces para todo entero positivo n existe un punto $\mathbf{x}_n$ en $\mathscr{E}$ tal que $|\mathbf{f}(\mathbf{x}_n)| > n$. Por tanto, lím $|\mathbf{f}(\mathbf{x}_n)| = \infty$. La sucesión $\{\mathbf{x}_n\}$ tiene un punto límite **c** en $\mathscr{E}$ y, por tanto, tiene una subsucesión $\{\mathbf{x}_{n_k}\}$ tal que lím $\mathbf{x}_{n_k} = \mathbf{c}$. Como **f** es continua sobre $\mathscr{E}$, lím $|\mathbf{f}(\mathbf{x}_{n_k})| = |\mathbf{f}(\mathbf{c})|$. Por otra parte, como lím $|\mathbf{f}(\mathbf{x}_n)| = \infty$, lím $|\mathbf{f}(\mathbf{x}_{n_k})| = \infty$. Esta contradicción demuestra que la hipótesis de que **f** no es acotada sobre $\mathscr{E}$ no puede verificarse. Y esto completa la prueba.

7.8 Teorema. *Si una función* **f** *de* R^m *en* R^n *es continua sobre un conjunto cerrado y acotado* $\mathscr{E}$, *entonces* $\mathbf{f}(\mathscr{E})$ *es un conjunto cerrado y acotado.*

PRUEBA. Que $\mathbf{f}(\mathscr{E})$ es un conjunto acotado se sigue del teorema 7.7. Para probar que $\mathbf{f}(\mathscr{E})$ es cerrado, probaremos que cualquier punto frontera **y** de $\mathbf{f}(\mathscr{E})$ pertenece a $\mathbf{f}(\mathscr{E})$. Como **y** es un punto frontera de $\mathbf{f}(\mathscr{E})$, cualquier vecindad de **y** contiene un punto de $\mathbf{f}(\mathscr{E})$. Luego para todo entero positivo n, existe un punto $\mathbf{y}_n \in \mathscr{S}\left(\mathbf{y}; \dfrac{1}{n}\right) \cap \mathbf{f}(\mathscr{E})$. Claramente, la sucesión $\{\mathbf{y}_n\}$ converge

a $\mathbf{y}$. Sea ahora $\{\mathbf{x}_n\}$ una sucesión de puntos en $\mathscr{E}$ tal que $\mathbf{f}(\mathbf{x}_n) = \mathbf{y}_n$. Como $\mathscr{E}$ es cerrado y acotado, $\{\mathbf{x}_n\}$ tiene un punto límite $\mathbf{x}$ en $\mathscr{E}$. Sea $\{\mathbf{x}_{n_k}\}$ una subsucesión de $\{\mathbf{x}_n\}$ que converge a $\mathbf{x}$. La subsucesión correspondiente $\{\mathbf{y}_{n_k}\}$ de $\{\mathbf{y}_n\}$ converge a $\mathbf{y}$. Como $\mathbf{f}$ es continua en $\mathbf{x}$ usamos el teorema 3.1 para obtener

$$\mathbf{y} = \lim_{k\to\infty} \mathbf{y}_{n_k} = \lim_{k\to\infty} \mathbf{f}(\mathbf{x}_{n_k}) = \mathbf{f}(\mathbf{x}).$$

Luego $\mathbf{y}\in\mathbf{f}(\mathscr{E})$, y $\mathbf{f}(\mathscr{E})$ es cerrado.

7.9 Teorema. *Si f es una función de R^m en R que es continua sobre un conjunto cerrado y acotado $\mathscr{E}$, entonces f tiene un valor máximo y un valor mínimo sobre $\mathscr{E}$.*

PRUEBA. Sea $b = \sup\{f(\mathbf{x}) \mid \mathbf{x}\in\mathscr{E}\}$. Para cualquier entero positivo n existe un punto $\mathbf{x}_n$ en $\mathscr{E}$ tal que

$$b - \frac{1}{n} < f(\mathbf{x}_n) \leqslant b.$$

Así pues, $\lim f(\mathbf{x}_n) = b$. La sucesión $\{\mathbf{x}_n\}$ tiene un punto límite $\mathbf{c}$ en $\mathscr{E}$ y, por tanto, tiene una subsucesión $\{\mathbf{x}_{n_k}\}$ tal que $\lim \mathbf{x}_{n_k} = \mathbf{c}$. Como f es continua sobre $\mathscr{E}$, $\lim f(\mathbf{x}_{n_k}) = f(\mathbf{c})$. Por otra parte, como $\lim f(\mathbf{x}_n) = b$, $\lim f(\mathbf{x}_{n_k}) = b$. Luego $f(\mathbf{c}) = b$ y b es el máximo valor de f sobre $\mathscr{E}$. Que f tiene un valor mínimo sobre $\mathscr{E}$, lo podemos probar de modo análogo.[1]

El teorema 7.9 nos muestra que una función f continua sobre un conjunto cerrado y acotado tiene un valor máximo y un valor mínimo sobre ese conjunto. Los valores máximo y mínimo pueden ocurrir solamente en puntos críticos de f que estén en el interior de $\mathscr{E}$ y en puntos de la frontera de $\mathscr{E}$. Así pues, podemos determinar los valores máximo y mínimo de f sobre $\mathscr{E}$ calculando los valores de f en esos puntos.

7.10 Ejemplo. Determínense los valores máximo y mínimo de la función f sobre el conjunto $\mathscr{E} = \{(x, y) \mid x^2+y^2 \leqslant 4\}$ cuando f está definida por

$$f(x, y) = \tfrac{3}{4}x^2+\tfrac{1}{3}y^2.$$

SOLUCIÓN. Como f es continua sobre el conjunto $\mathscr{E}$ que es cerrado y acotado, el teorema 7.9 asegura la existencia de un valor máximo y un valor mínimo de f sobre $\mathscr{E}$. Las derivadas parciales de f son

$$D_1 f(x, y) = \tfrac{3}{2}x \quad \text{y} \quad D_2 f(x, y) = \tfrac{2}{3}y,$$

estas derivadas parciales existen en todos los puntos y son, ambas, cero solamente en $(0, 0)$. Así pues, el único punto crítico de f es $(0, 0)$ y

[1] Lo habitual es pasar a considerar $-f$. [N. del T.]

$f(0, 0) = 0$. En la frontera de $\mathscr{E}$, $\{(x, y) \mid x^2+y^2 = 4\}$, los valores de f están dados por

$$f(x, y) = \tfrac{5}{12}x^2+\tfrac{4}{3} \quad \text{donde} \quad x\in[-2, 2].$$

Por tanto, sobre la frontera de $\mathscr{E}$ la función f toma todos los valores del intervalo cerrado $[\frac{3}{4}, 3]$. Así pues, el valor máximo de f sobre $\mathscr{E}$ es 3 y el valor mínimo es 0. El diagrama de las curvas de nivel y la gráfica de f aparecen dibujados en la figura 3.

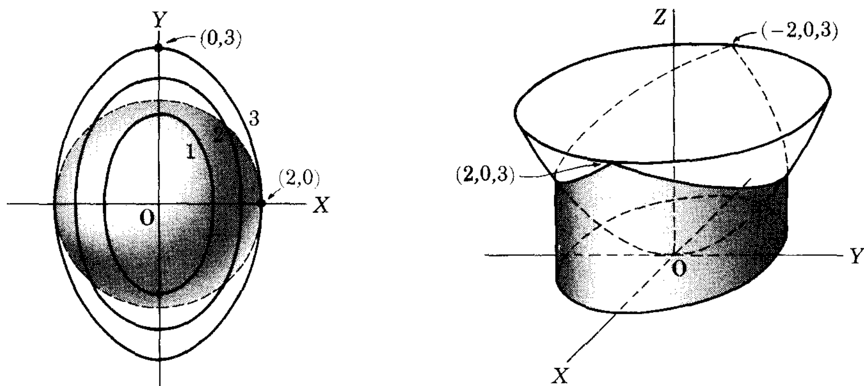


FIGURA 3

Problemas

1. Supongamos que $\{\mathbf{x}_n\}$ es una sucesión acotada de puntos en R^m. Pruébese que $\lim \mathbf{x}_n = \mathbf{c}$ si y sólo si $\mathbf{c}$ es el único punto límite de $\{\mathbf{x}_n\}$.

2. Sea $\{\mathbf{x}_n \mid n = 0, 1, \dots\}$ una sucesión de puntos en R^m con la propiedad de que para algún $\lambda\in\langle 0, 1\rangle$

$$|\mathbf{x}_n-\mathbf{x}_{n-1}| \leqslant \lambda|\mathbf{x}_{n-1}-\mathbf{x}_{n-2}|$$

para todo $n \geqslant 2$.

a) Pruébese que para $n \geqslant 0$ y $k \geqslant 1$

$$|\mathbf{x}_{n+k}-\mathbf{x}_n| \leqslant \frac{\lambda^n}{1-\lambda}|\mathbf{x}_1-\mathbf{x}_0|.$$

b) Pruébese que $\{\mathbf{x}_n\}$ es acotada.
c) Pruébese que $\{\mathbf{x}_n\}$ tiene un punto límite único y, por tanto, la sucesión converge.

3. Determínense los valores máximo y mínimo de las siguientes

funciones f sobre el conjunto dado $\mathscr{E}$. Dibújense el diagrama de curvas de nivel y la gráfica en cada uno de los casos.

a) $f(x, y) = xy - y$ $\qquad \mathscr{E} = \{(x, y) \mid |x| \leqslant 3,\ |y| \leqslant 2\}$
b) $f(x, y) = \frac{1}{3}x^2 - \frac{3}{16}y^2$ $\qquad \mathscr{E} = \{(x, y) \mid |x| \leqslant 3,\ |y| \leqslant 3\}$
c) $f(x, y) = x^2y^2$ $\qquad \mathscr{E} = \{(x, y) \mid x^2 + y^2 \leqslant 9\}$
d) $f(x, y) = 5 - x^2 - y^2$ $\qquad \mathscr{E} = \{(x, y) \mid |x| \leqslant 4,\ |y| \leqslant 4\}$
e) $f(x, y) = 1 + x^2 - y^2$ $\qquad \mathscr{E} = \{(x, y) \mid x^2 + y^2 \leqslant 4\}$
f) $f(x, y) = 3 - \sqrt{x^2 + y^2}$ $\qquad \mathscr{E} = \{(x, y) \mid 4x^2 + 9y^2 \leqslant 36\}$.

8. SUCESIONES DE FUNCIONES

Consideraremos ahora sucesiones cuyos términos son funciones.

8.1 Definición. *Una* ***sucesión de funciones*** *de* R^p *en* R^m *es una función cuyo dominio es el conjunto de los enteros positivos y cuyo rango es un conjunto de funciones de* R^p *en* R^m. *A una sucesión de funciones de* R^p *en* R^m *la representamos por* $\{\mathbf{f}_n\}$.

Para cada punto x del dominio de todos los términos de la sucesión de funciones $\{\mathbf{f}_n\}$, hay una sucesión de puntos $\{\mathbf{f}_n(\mathbf{x})\}$. Si $\{\mathbf{f}_n(\mathbf{x})\}$ converge para cada punto $\mathbf{x}$ de un conjunto $\mathscr{E}$ y hacemos $\mathbf{f}(\mathbf{x}) = \lim\limits_{n\to\infty} \mathbf{f}_n(\mathbf{x})$, entonces decimos que $\{\mathbf{f}_n\}$ *converge puntualmente* a $\mathbf{f}$ sobre $\mathscr{E}$.

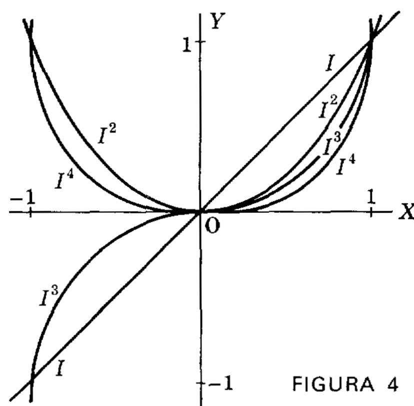


FIGURA 4

Consideremos, por ejemplo, la sucesión $\{I^n\}$ de funciones reales de variable real donde I es la función real identidad (figura 4). Para cualquier $x \in \mathrm{R}$, $\{x^n\}$ es una sucesión de números reales cada término de la cual es el valor de la función en x del correspondiente término de $\{I^n\}$. Como $\{x^n\}$ converge a 0 si $x \in \langle -1, 1\rangle$, converge a 1 si $x = 1$, y diverge si

$$x \in \langle -\infty, -1] \cup \langle 1, \infty\rangle,$$

la sucesión $\{I^n\}$ converge puntualmente sobre $\langle -1, 1]$ a la función f donde $f(x) = 0$ si $x \in \langle -1, 1\rangle$ y $f(1) = 1$.

Nótese que, aunque $\{I^n\}$ es una sucesión de funciones continuas, la función límite f no es continua.

Definimos ahora otro tipo de convergencia para sucesiones de funciones —convergencia uniforme— y probaremos que el límite de una sucesión uniformemente convergente de funciones continuas es continua.

8.2 Definición. *Una sucesión de funciones* $\{\mathbf{f}_n\}$ ***converge uniformemente*** *a* $\mathbf{f}$ *sobre un conjunto* $\mathscr{E}$ *si para cada* $\varepsilon > 0$ *existe un número* $\mathcal{N}$ *tal que para todo* $\mathbf{x} \in \mathscr{E}$

$$|\mathbf{f}_n(\mathbf{x}) - \mathbf{f}(\mathbf{x})| < \varepsilon \quad \textit{siempre que } n > N.$$

Lo que es importante observar respecto a la convergencia uniforme es que el número N depende solamente de ε; es independiente del punto $\mathbf{x}$ de $\mathscr{E}$. Por otra parte, si $\{\mathbf{f}_n\}$ converge puntualmente a $\mathbf{f}$ sobre $\mathscr{E}$, entonces para cualquier $\mathbf{x} \in \mathscr{E}$ y para cualquier número positivo ε existe un número $N(\mathbf{x})$, que depende en general tanto de ε como de $\mathbf{x}$, tal que

$$|\mathbf{f}_n(\mathbf{x}) - \mathbf{f}(\mathbf{x})| < \varepsilon \quad \text{siempre que } n > N(\mathbf{x}).$$

Es claro que si $\{\mathbf{f}_n\}$ converge uniformemente a $\mathbf{f}$ sobre $\mathscr{E}$, entonces la sucesión es puntualmente convergente a $\mathbf{f}$ sobre $\mathscr{E}$. Pero la convergencia puntual de $\{\mathbf{f}_n\}$ a $\mathbf{f}$ sobre $\mathscr{E}$ no implica la convergencia uniforme de la sucesión a $\mathbf{f}$ sobre $\mathscr{E}$. En realidad, podemos ver que una sucesión puntualmente convergente $\{\mathbf{f}_n\}$ sobre $\mathscr{E}$ es uniformemente convergente sobre $\mathscr{E}$ si y sólo si para cada $\varepsilon > 0$ existe un conjunto de números correspondientes $\{N(\mathbf{x}) \mid \mathbf{x} \in \mathscr{E}\}$ que es superiormente acotado. Si $\{\mathbf{f}_n\}$ es convergente puntualmente a $\mathbf{f}$ sobre $\mathscr{E}$ y N es una cota superior de $\{N(\mathbf{x}) \mid \mathbf{x} \in \mathscr{E}\}$, entonces para todo $\mathbf{x} \in \mathscr{E}$

$$|\mathbf{f}_n(\mathbf{x}) - \mathbf{f}(\mathbf{x})| < \varepsilon \quad \text{siempre que } n > N.$$

Así pues, $\{\mathbf{f}_n\}$ es uniformemente convergente a $\mathbf{f}$ sobre $\mathscr{E}$. Además, si $\{\mathbf{f}_n\}$ es uniformemente convergente a $\mathbf{f}$ sobre $\mathscr{E}$ y N corresponde a $\varepsilon > 0$, entonces para cada $\mathbf{x} \in \mathscr{E}$ podemos tomar $N(\mathbf{x}) = N$ y el conjunto $\{N(\mathbf{x}) \mid \mathbf{x} \in \mathscr{E}\}$ quedará superiormente acotado por N.

8.3 Ejemplo. Pruébese que $\{I^n\}$ es uniformemente convergente sobre $[-a, a]$ donde $0 < a < 1$.

SOLUCIÓN. Para cada $x \in [-a, a]$, $\lim_{n \to \infty} x^n = 0$ e $|I^n(x) - 0| = |x|^n \leqslant a^n$. Para cualquier $\varepsilon > 0$, $a^n < \varepsilon$ si $n > \dfrac{\ln \varepsilon}{\ln a}$. Así pues, si $N = \dfrac{\ln \varepsilon}{\ln a}$, se tiene para cualquier $x \in [-a, a]$,

$$|I^n(x) - 0| < \varepsilon \quad \text{siempre que } n > N.$$

Lo que demuestra que $\{I^n\}$ converge uniformemente a 0 sobre $[-a, a]$.

8.4 Ejemplo. Pruébese que $\{I^n\}$ no es uniformemente convergente sobre $\langle -1, 1\rangle$.

SOLUCIÓN. Tómese un ε cualquiera tal que $0 < \varepsilon < 1$. Para cada $x \in \langle -1, 1\rangle$, $\lim x^n = 0$ y por tanto, existe un número $N(x)$ tal que

$$|I^n(x) - 0| = |x|^n < \varepsilon \quad \text{siempre que } n > N(x).$$

Como $|x|^n < \varepsilon$ si y sólo si $n > \dfrac{\ln \varepsilon}{\ln |x|}$ y como $\lim\limits_{x \to 1^-} \dfrac{\ln \varepsilon}{\ln x} = \infty$, el conjunto $\{N(x) \mid x \in \langle -1, 1\rangle\}$ no puede estar acotado superiormente. Así pues, no puede haber ningún número N tal que para todo $x \in \langle -1, 1\rangle$

$$|I^n(x) - 0| < \varepsilon \quad \text{siempre que } n > N.$$

Ahora probaremos que una sucesión uniformemente convergente de funciones continuas converge a una función continua.

8.5 Teorema. *Si la sucesión $\{\mathbf{f}_n\}$ converge uniformemente a $\mathbf{f}$ sobre el conjunto $\mathscr{E}$ y cada uno de los términos $\mathbf{f}_n$ es continuo sobre $\mathscr{E}$, entonces $\mathbf{f}$ es continua sobre $\mathscr{E}$.*

PRUEBA. Sea $\mathbf{x}_0$ un punto cualquiera en $\mathscr{E}$ y tomemos un $\varepsilon > 0$ cualquiera. Deseamos probar que hay un número $\delta > 0$ tal que

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)| < \varepsilon \quad \text{siempre que } \mathbf{x} \in \mathscr{E} \cap \mathscr{S}(\mathbf{x}_0; \delta).$$

La convergencia uniforme de $\{\mathbf{f}_n\}$ a $\mathbf{f}$ sobre $\mathscr{E}$ implica que para algún entero positivo n

$$|\mathbf{f}_n(\mathbf{x}) - \mathbf{f}(\mathbf{x})| < \frac{\varepsilon}{3} \quad \text{para todo } \mathbf{x} \in \mathscr{E}.$$

Como $\mathbf{f}_n$ es continua en $\mathbf{x}_0$, existe una $\delta > 0$ tal que

$$|\mathbf{f}_n(\mathbf{x}) - \mathbf{f}_n(\mathbf{x}_0)| < \frac{\varepsilon}{3} \quad \text{siempre que } \mathbf{x} \in \mathscr{E} \cap \mathscr{S}(\mathbf{x}_0; \delta).$$

Por tanto, para todo $\mathbf{x} \in \mathscr{E} \cap \mathscr{S}(\mathbf{x}_0; \delta)$

$$\begin{aligned} |\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)| &\leqslant |\mathbf{f}(\mathbf{x}) - \mathbf{f}_n(\mathbf{x})| + |\mathbf{f}_n(\mathbf{x}) - \mathbf{f}_n(\mathbf{x}_0)| + |\mathbf{f}_n(\mathbf{x}_0) - \mathbf{f}(\mathbf{x}_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Lo que completa la prueba.

Aunque la convergencia uniforme es suficiente para asegurar que el límite de una sucesión de funciones continuas es continua, no es una condición necesaria para este resultado. Se ve esto en el siguiente ejemplo de sucesión no uniformemente convergente de funciones continuas cuyo límite es una función continua. Sea

$$f_1(x) = 1 \quad \text{si} \quad x \in [0, 1]$$

y para $n \geqslant 2$ (figura 5) sea

8.6
$$f_n(x) = \begin{cases} nx & \text{si} \quad x \in \left[0, \dfrac{1}{n}\right] \\ 2 - nx & \text{si} \quad x \in \left\langle \dfrac{1}{n}, \dfrac{2}{n} \right] \\ 0 & \text{si} \quad x \in \left\langle \dfrac{2}{n}, 1 \right] \end{cases}$$

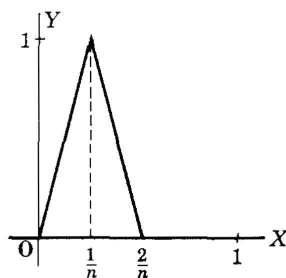


FIGURA 5

Para cada $x \in [0, 1]$, $\lim\limits_{n \to \infty} f_n(x) = 0$ ya que $f_n(x) = 0$ para $n > \dfrac{2}{x}$ si $x \neq 0$ y $f_n(0) = 0$ para todo $n > 1$. Así pues, la función límite f es la función cero sobre $[0, 1]$ que es ciertamente continua. Sin embargo, la convergencia no es uniforme ya que para cualquier n si $x = \dfrac{1}{n}$ entonces $|f_n(x) - f(x)| = 1$.

Damos ahora algunos resultados respecto a la integración y diferenciación de los términos de una sucesión de funciones reales de variable real.

8.7 Teorema. *Si la sucesión* $\{f_n\}$ *converge uniformemente a* f *sobre el intervalo cerrado* $[a, b]$ *y si cada uno de los términos* f_n *es integrable*

sobre $[a, b]$, *entonces* f *es integrable sobre* $[a, b]$. *Además, si*

$$g_n(x) = \int_a^x f_n \quad y \quad g(x) = \int_a^x f, \qquad x \in [a, b],$$

entonces $\{g_n\}$ *converge uniformemente a* g *sobre* $[a, b]$.

PRUEBA. Probaremos primero que f es integrable sobre $[a, b]$. Tómese $\varepsilon > 0$. Como $\{f_n\}$ converge uniformemente a f sobre $[a, b]$, para alguna n

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3(b-a)} \quad \text{para todo } x \in [a, b].$$

Así pues, para cualquier partición P de $[a, b]$,

$$m_i(f) \geqslant m_i(f_n) - \frac{\varepsilon}{3(b-a)} \quad \text{y} \quad M_i(f) \leqslant M_i(f_n) + \frac{\varepsilon}{3(b-a)}$$

y, por tanto

$$L(f, P) \geqslant L(f_n, P) - \frac{\varepsilon}{3} \quad \text{y} \quad U(f, P) \leqslant U(f_n, P) + \frac{\varepsilon}{3}.$$

Como f_n es integrable sobre $[a, b]$, $U(f_n, P) - L(f_n, P) < \frac{\varepsilon}{3}$ para alguna partición P de $[a, b]$. Para esta partición tenemos

$$U(f, P) - L(f, P) < \varepsilon.$$

Lo que prueba que f es integrable sobre $[a, b]$. Probamos ahora que $\{g_n\}$ converge uniformemente a g sobre $[a, b]$; es decir, para cualquier $\varepsilon > 0$ existe un número N tal que para todo $x \in [a, b]$

$$\left| \int_a^x f - \int_a^x f_n \right| < \varepsilon \quad \text{siempre que } n > N.$$

Como $\{f_n\}$ converge uniformemente a f sobre $[a, b]$, existe un número N tal que para todo $t \in [a, b]$

$$|f_n(t) - f(t)| < \frac{\varepsilon}{2(b-a)} \quad \text{siempre que } n > N.$$

Por tanto, para todo $n > N$ y para toda $x \in [a, b]$

$$\left| \int_a^x f - \int_a^x f_n \right| = \left| \int_a^x (f - f_n) \right| \leqslant \frac{\varepsilon}{2(b-a)} (x-a) < \varepsilon.$$

Esto completa la prueba.

Nota. Recuérdese que si f_n es continua sobre $[a, b]$, entonces f_n es integrable sobre $[a, b]$.

El teorema 8.7 demuestra que la convergencia uniforme de una sucesión de funciones integrables $\{f_n\}$ a f sobre $[a, b]$ implica que

8.8
$$\int_a^b f = \int_a^b \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_a^b f_n .$$

Sin embargo 8.8 puede ser cierto para una sucesión no uniformemente convergente. Por ejemplo, si $\{f_n\}$ es la sucesión definida en 8.6, entonces

$$\lim_{n\to\infty} \int_0^1 f_n = \lim_{n\to\infty} \frac{1}{n} = 0 = \int_0^1 f.$$

Volviendo ahora a una consideración de diferenciación de los términos de una sucesión, nos encontramos con que si $\{f_n\}$ es una sucesión uniformemente convergente de funciones diferenciables con la función f como límite, no es necesariamente cierto que $f' = \lim f_n'$.

Por ejemplo, sea $f_n(x) = \frac{1}{n}$ sen nx. Entonces la sucesión $\{f_n\}$ converge a 0 sobre R. Además, para todo $x \in R$ y para cualquier $\varepsilon > 0$

$$\left|\frac{1}{n} \text{ sen } nx - 0\right| = \frac{1}{n} |\text{sen } nx| \leqslant \frac{1}{n} < \varepsilon \quad \text{siempre que } n > \frac{1}{\varepsilon}.$$

Así pues, $\{f_n\}$ converge uniformemente a $f = 0$ sobre R y $f' = 0$. Como $f_n'(x) = \cos nx$, la sucesión de derivadas $\{f_n'\}$ no converge sobre R; es decir, $\left\{\cos n \frac{\pi}{4}\right\}$ no converge. Por tanto, $f' \neq \lim f_n'$.

El siguiente teorema nos da una condición suficiente para que f' sea el límite de la sucesión $\{f_n'\}$.

8.9 Teorema. *Si* $\{f_n\}$ *converge a* f *sobre* $[a, b]$, *si cada una de las derivadas* f_n' *es continua sobre* $[a, b]$, *y si* $\{f_n'\}$ *converge uniformemente sobre* $[a, b]$, *entonces* $f' = \lim f_n'$ *sobre* $[a, b]$.

PRUEBA. Sea $g = \lim f_n'$. Deseamos probar que $g = f'$ sobre $[a, b]$. Sea $x \in [a, b]$. Como $\{f_n'\}$ converge uniformemente a g sobre $[a, b]$, tenemos

$$\int_a^x g = \lim_{n\to\infty} \int_a^x f_n' = \lim_{n\to\infty} [f_n(x) - f_n(a)] = f(x) - f(a).$$

Tomando derivadas, obtenemos $g(x) = f'(x)$ y, por tanto, $\lim f_n' = f'$ sobre $[a, b]$.

Problemas

1. Determínese el conjunto sobre el que $\{f_n\}$ converge si $f_n(x)$ es

$a)\ \dfrac{x^n}{n}$ $\quad b)\ \dfrac{2n+x}{n-3}$ $\quad c)\ \dfrac{2}{nx+1}$

$d)\ nxe^{-nx}$ $\quad e)\ \dfrac{x^n}{n!}$ $\quad f)\ \dfrac{xn^2+3}{n+1}$.

2. Si $f_n(x) = \dfrac{x^n}{n}$ pruébese que $\{f_n\}$ es uniformemente convergente sobre $[-1, 1]$.

3. Si $f_n(x) = x$ y $g_n(x) = x + \dfrac{1}{n}$, demuéstrese que $\{f_n\}$ y $\{g_n\}$ son uniformemente convergentes sobre $\langle -\infty, \infty\rangle$.

4. Si $f_n(x) = \dfrac{2n+x}{n+3}$ pruébese que $\{f_n\}$ es uniformemente convergente sobre cualquier intervalo $[-a, a]$ donde $a > 0$.

5. Si $f_n(x) = nxe^{-nx}$ pruébese que $\{f_n\}$ no es uniformemente convergente sobre $[0, 1]$. *Sugerencia :* considérese el valor máximo de f_n sobre $[0, 1]$.

6. Pruébese que $\{\mathbf{f}_n\}$ es uniformemente convergente a $\mathbf{f}$ sobre el conjunto $\mathscr{E}$ si y sólo si $\lim_{n\to\infty} \sup_{\mathbf{x}\in\mathscr{E}} |\mathbf{f}(\mathbf{x}) - \mathbf{f}_n(\mathbf{x})| = 0$.

7. Sea $\{\mathbf{f}_n\}$ una sucesión de funciones de R^p en R^m y $\mathbf{f}_n = (f_n^1, \ldots, f_n^p)$ donde f_n^k es la función componente k-ésima de $\mathbf{f}_n$. Demuéstrese que $\{\mathbf{f}_n\}$ es uniformemente convergente a $\mathbf{f} = f^1, \ldots, f^p$ sobre $\mathscr{E}$ si y sólo si cada una de las sucesiones de las componentes $\{f_n^k\}$ es uniformemente convergente a f^k sobre $\mathscr{E}$.

8. Si $\{f_n\}$, $\{g_n\}$ y $\{h_n\}$ son sucesiones de funciones de R^p en R, si $\{f_n\}$ y $\{h_n\}$ son uniformemente convergentes a f en $\mathscr{E}$, y si para toda n y para toda $\mathbf{x}\in\mathscr{E}$

$$f_n(\mathbf{x}) \leqslant g_n(\mathbf{x}) \leqslant h_n(\mathbf{x})$$

pruébese que $\{g_n\}$ converge uniformemente a f sobre $\mathscr{E}$.

9. Si $f_n(x) = e^{-n^2x^2}$, pruébese que $\{f_n'\}$ no es uniformemente convergente sobre $[0, 1]$. Pruébese que, sin embargo, $f' = \lim f_n'$ donde $f = \lim f_n$.

10. Si $f_n(x) = (\text{sen}\ \pi x)^n$, ¿es $\{f_n\}$ uniformemente convergente sobre $[0, 1]$? ¿Es $\{f_n'\}$ uniformemente convergente sobre $[0, 1]$?

11. Si $f_n(x) = \dfrac{nx}{1+n^2x^2}$, ¿es $\{f_n\}$ uniformemente convergente sobre $[a, \infty]$ donde $a > 0$? ¿Es $\{f_n\}$ uniformemente convergente sobre $\langle 0, \infty \rangle$?

12. Si $\{\mathbf{f}_n\}$ y $\{\mathbf{g}_n\}$ son uniformemente convergentes sobre un conjunto $\mathscr{E}$ pruébese que $\{\mathbf{f}_n + \mathbf{g}_n\}$ es uniformemente convergente sobre $\mathscr{E}$.

13. Proporciónese un ejemplo de sucesiones $\{f_n\}$ y $\{g_n\}$ que son uniformemente convergentes sobre un conjunto $\mathscr{E}$, pero tales que $\{f_n g_n\}$ no es uniformemente convergente sobre $\mathscr{E}$.

9. RESUMEN

En este capítulo hemos considerado límites de sucesiones de puntos. Como la convergencia de una sucesión de puntos depende de la convergencia de sus sucesiones componentes, dirigimos nuestra mayor atención a sucesiones de números reales. Vimos que las sucesiones de números reales pueden convergir, divergir a $\pm\infty$ u oscilar, y desarrollamos métodos para determinar el comportamiento de sucesiones específicas.

En el próximo capítulo estudiaremos las series, que son sucesiones formadas de un modo particular. La mayor parte del material que hemos discutido en este capítulo es prerrequisito indispensable para un estudio de la convergencia de las series.

Problemas de repaso

1. Si $s_n = \sum_{k=0}^{n-1} r^k$, discútase la convergencia de $\{s_n\}$.

2. Determínese lím s_n cuando s_n es

a) $\dfrac{3n^2 - 5n + 2}{4n^2 + 7}$ *b*) $\dfrac{-3n^2 + 5}{n + 3}$

c) $\ln \dfrac{n+2}{n^2 + 5n}$ *d*) $\dfrac{\ln n + 2}{n - 3}$

e) $\dfrac{n^4 + e^n}{n - 5^n}$ *f*) $\dfrac{n!}{3^n + n^2}$.

3. Si lím $x_n = \infty$ y lím $y_n = 0$, pruébese, por medio de ejemplos, que todos los casos que siguen pueden ocurrir

a) lím $x_n y_n = \infty$ *b*) lím $x_n y_n = -\infty$
c) lím $x_n y_n = b$ (un número real) *d*) lím $x_n y_n \neq b, \pm\infty$.

4. Pruébese que si $s_n < 0$ y lím $s_n = 0$, entonces lím $\frac{1}{s_n} = -\infty$.

5. Pruébese lo siguiente:

a) Si $s_n > 0$ y lím $\frac{s_{n+1}}{s_n} > 1$, entonces lím $s_n = \infty$

b) Si $s_n < 0$ y lím $\frac{s_{n+1}}{s_n} > 1$, entonces lím $s_n = -\infty$

c) Si $s_n = b \neq 0$, entonces lím $\frac{s_{n+1}}{s_n} = 1$.

6. Proporciónense ejemplos de sucesiones $\{s_n\}$ tales que lím $\frac{s_{n+1}}{s_n} = 1$ y

a) lím $s_n = b \neq 0$ *b*) lím $s_n = 0$ *c*) lím $s_n = \infty$.

7. Proporciónense ejemplos de sucesiones $\{s_n\}$ tales que lím $\frac{s_{n+1}}{s_n} = -1$ y

a) $\{s_n\}$ converge a 0 *b*) $\{s_n\}$ oscila.

8. Definimos una sucesión de Cauchy de números reales como sigue: una sucesión $\{s_n\}$ se llama *sucesión de Cauchy* si, para cualquier $\varepsilon > 0$, existe un número N tal que $|s_n - s_m| < \varepsilon$ siempre que $n, m > N$.

a) Pruébese que toda sucesión convergente es una sucesión de Cauchy.
b) Pruébese que una sucesión de Cauchy es acotada y, por tanto, tiene un punto límite finito.
c) Pruébese que una sucesión de Cauchy converge.

9. Pruébese que si $s_n > 0$

$$\underline{\text{lím}}\ \frac{s_{n+1}}{s_n} \leqslant \underline{\text{lím}}\ \sqrt[n]{s_n} \leqslant \overline{\text{lím}}\ \sqrt[n]{s_n} \leqslant \overline{\text{lím}}\ \frac{s_{n+1}}{s_n}.$$





Series

1. INTRODUCCIÓN

El teorema de Taylor para funciones reales de una variable real, nos dice: *si f tiene derivadas continuas hasta las de orden n-ésimo inclusive sobre el intervalo $\mathscr{I}$ y $x_0 \in \mathscr{I}$, entonces para todo $x \in \mathscr{I}$*

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_n(x)$$

donde

$$R_n(x) = \frac{1}{(n-1)!}\int_{x_0}^{x} f^{(n)}(t)\,(x-t)^{n-1}\,dt\,.$$

Si $x \neq x_0$, usando el primer teorema para el valor medio para integrales,

podemos escribir el residuo $R_n(x)$ de la siguiente forma (forma de Lagrange):

$$R_n(x) = \frac{f^{(n)}(c_n)}{n!}(x-x_0)^n \text{ para algún } c_n \text{ entre } x \text{ y } x_0 .$$

Si $x = x_0$, entonces $R_n(x_0)$ es cero y la fórmula de Taylor simplemente afirma que $f(x_0) = f(x_0)$. En lugar de señalar excepciones sin importancia supondremos que el residuo puede escribirse en la forma de Lagrange para cualquier valor de x y convendremos que es cero cuando $x = x_0$.

Por ejemplo, si f es la función exponencial y x_0 es 0; entonces, para cualquier número real x

1.1
$$e^x = \sum_{k=0}^{n-1} \frac{x^k}{k!} + R_n(x)$$

donde $R_n(x) = \dfrac{e^{c_n}}{n!} x^n$ para algún c_n entre x y 0. Como la función exponencial tiene derivadas continuas de todos los órdenes, la n en 1.1 puede ser cualquier entero positivo. Para cada entero positivo n, sea $s_n(x) = \sum_{k=0}^{n-1} \dfrac{x^k}{k!}$. Entonces, para cada número real x, tenemos dos sucesiones, $\{s_n(x)\}$ y $\{R_n(x)\}$, con

1.2
$$e^x = s_n(x) + R_n(x).$$

Probaremos ahora que $\lim\limits_{n\to\infty} R_n(x) = 0$ y, por tanto, $e^x = \lim\limits_{n\to\infty} s_n(x)$. Si $x > 0$, entonces

$$0 < R_n(x) = \frac{e^{c_n}}{n!} x^n < e^x \frac{x^n}{n!} .$$

Como $\lim\limits_{n\to\infty} \dfrac{x^n}{n!} = 0$, $\lim\limits_{n\to\infty} R_n(x) = 0$. Si $x < 0$, entonces

$$0 < |R_n(x)| < \frac{|x|^n}{n!}$$

y, por tanto, $\lim\limits_{n\to\infty} R_n(x) = 0$. Por tanto. para cualquier número real x,

1.3
$$e^x = \lim_{n\to\infty} s_n(x) = \lim_{n\to\infty} \sum_{k=0}^{n-1} \frac{x^k}{k!} .$$

En 1.3 el número e^x está expresado como el límite de una suma de números; a medida que sumamos más y más términos obtenemos números más y más próximos a e^x. Así pues, consideramos a e^x como la suma

infinita $\sum_{k=0}^{\infty} \frac{x^k}{k!}$. Desde luego, ni sumamos ni podemos sumar realmente infinitos números; lo que hacemos es tomar el límite de las sumas finitas $s_n(x)$. La sucesión $\{s_n(x)\}$ se llama serie infinita o simplemente serie. Si $\lim_{n\to\infty} s_n(x)$ existe, como, por ejemplo, ocurre en el caso anterior, entonces decimos que la serie converge.

Las series son útiles tanto en el cálculo como en el estudio de las propiedades de las funciones. Como las series están definidas en términos de sumas, es de esperarse que tengan propiedades análogas a las de las sumas. Veremos que, ciertamente, las series convergentes poseen la mayoría de las propiedades algebraicas de las sumas. Para algunas otras necesitamos que la convergencia sea uniforme. Son todas estas cuestiones las que discutiremos ampliamente en este capítulo.

2. SERIES

2.1 Definición. *Sea* $\{\mathbf{a}_k\}$ *una sucesión de puntos en* R^m *y* $\mathbf{s}_n = \sum_{k=1}^{n} \mathbf{a}_k$. *A la sucesión* $\{\mathbf{s}_n\}$ *se le llama* ***serie*** *y a los términos de* $\{\mathbf{a}_k\}$ *se les llama términos de la serie.*

Comúnmente denotamos las series $\{\mathbf{s}_n\}$ por $\sum_{k=1}^{\infty} \mathbf{a}_k$ o, simplemente, por $\Sigma\mathbf{a}_k$. Si la sucesión $\{\mathbf{s}_n\}$ converge al punto $\mathbf{a}$, entonces decimos que la serie $\sum_{k=1}^{\infty} \mathbf{a}_k$ tiene la *suma* $\mathbf{a}$ o que $\sum_{k=1}^{\infty} \mathbf{a}_k$ *converge* a $\mathbf{a}$. La suma $\mathbf{s}_n$ de los primeros n términos de la serie se llama algunas veces *suma parcial.* Así pues, nuestra definición nos dice que una serie converge si y sólo si la sucesión de sumas parciales converge. Como una sucesión puede convergir o no, una serie puede tener una suma o puede ser que no la tenga.

Si la serie $\sum_{k=1}^{\infty} \mathbf{a}_k$ tiene la suma $\mathbf{a}$, entonces $\mathbf{a} = \lim \mathbf{s}_n$ donde $\mathbf{s}_n = \sum_{k=1}^{n} \mathbf{a}_k$. De donde $\mathbf{a} = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbf{a}_k$. Es práctica constante escribir $\mathbf{a} = \sum_{k=1}^{\infty} \mathbf{a}_k$. Así pues, la notación $\sum_{k=1}^{\infty} \mathbf{a}_k$ se usa tanto para representar una serie como para representar su suma. Sin embargo, este uso dual del símbolo $\sum_{k=1}^{\infty} \mathbf{a}_k$ no debe causar ninguna confusión.

Como una serie es una sucesión, la teoría de las sucesiones desarrollada en el capítulo anterior se aplica a las series. Una sucesión de puntos converge si y sólo si las sucesiones componentes convergen. Así pues, una

serie de puntos converge si y sólo si las series componentes convergen. Por ejemplo, si $\{\mathbf{a}_k\}$ es una sucesión de puntos (x_k, y_k) en R^2, entonces

$$\lim_{n\to\infty} \sum_{k=1}^{n} \mathbf{a}_k = \lim_{n\to\infty} \sum_{k=1}^{n} (x_k, y_k) = \lim_{n\to\infty} \left(\sum_{k=1}^{n} x_k, \sum_{k=1}^{n} y_k \right)$$

y $\sum_{k=1}^{\infty} \mathbf{a}_k$ converge si y sólo si $\sum_{k=1}^{\infty} x_k$ y $\sum_{k=1}^{\infty} y_k$ convergen. Por tanto, podemos restringir nuestra atención a las series de números reales.

Consideremos las series geométricas $\sum_{k=1}^{\infty} x^{k-1}$ o $\sum_{k=0}^{\infty} x^k$.

Nota. Usualmente, la serie geométrica $\sum_{k=1}^{\infty} x^{k-1}$ se escribe en la forma $\sum_{k=0}^{\infty} x^k$. En general, cuando ello resulta más conveniente, podemos escribir una serie en la forma $\sum_{k=p}^{\infty} a_k$ donde p es un entero cualquiera. La notación $\sum_{k=p}^{\infty} a_k$ significa lo mismo que $\sum_{k=1}^{\infty} b_k$ donde $b_k = a_{k+p-1}$ y, con esta notación $s_n = \sum_{k=1}^{n} b_k = \sum_{k=p}^{p+n-1} a_k$.

En la investigación de la convergencia de las series geométricas $\sum_{k=0}^{\infty} x^k$ consideramos dos casos: $x = 1$ y $x \neq 1$. Si $x = 1$, entonces

$$s_n = \sum_{k=0}^{n-1} 1 = n$$

y, por tanto, $\{s_n\}$ diverge. Si $x \neq 1$, entonces

$$s_n - xs_n = \sum_{k=0}^{n-1} x^k - \sum_{k=1}^{n} x^k = 1 - x^n$$

y

$$s_n = \frac{1-x^n}{1-x}.$$

Como la sucesión $\{x^n\}$ converge a cero si $|x| < 1$ y diverge si $|x| > 1$ o $x = -1$, $\{s_n\}$ converge a $\dfrac{1}{1-x}$ si $|x| < 1$ y diverge si $|x| > 1$ o $x = -1$.

Así pues, hemos probado lo siguiente: *la serie geométrica* $\sum_{k=0}^{\infty} x^k$ *tiene la suma* $\dfrac{1}{1-x}$ *si* $|x| < 1$ *y* $\sum_{k=0}^{\infty} x^k$ *diverge si* $|x| \geqslant 1$.

2.2 Ejemplo. Pruébese que $\sum_{k=0}^{\infty} \frac{1}{2^k} = 2.$

Solución. $\sum_{k=0}^{\infty} \frac{1}{2^k}$ es la serie geométrica $\sum_{k=0}^{\infty} x^k$ con $x = \frac{1}{2}$. Como $|x| < 1$, la serie converge y

$$\sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{1-\frac{1}{2}} = 2\,.$$

Daremos ahora algunas propiedades de las series de números reales que se corresponden con propiedades sencillas de las sumas finitas de los números reales. Usando la teoría de las sucesiones, obtenemos estas propiedades de las series partiendo de las propiedades correspondientes de las sumas finitas.

2.3 Teorema. *Si* $\sum_{k=1}^{\infty} a_k$ *y* $\sum_{k=1}^{\infty} b_k$ *son series convergentes con sumas* a *y* b, *respectivamente, y si* c *es un número real, entonces*

1) $\sum_{k=1}^{\infty} (a_k+b_k)$ *converge a* $a+b$; *es decir,* $\sum_{k=1}^{\infty} (a_k+b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k$

2) $\sum_{k=1}^{\infty} (a_k-b_k)$ *converge a* $a-b$; *es decir,* $\sum_{k=1}^{\infty} (a_k-b_k) = \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} b_k$

3) $\sum_{k=1}^{\infty} ca_k$ *converge a* ca; *es decir,* $\sum_{k=1}^{\infty} ca_k = c\sum_{k=1}^{\infty} a_k$.

Prueba. Solamente probaremos 1). Las pruebas de las otras partes son análogas. Usando el corolario 3.2, pág. 458,

$$\begin{aligned}
\lim_{n\to\infty} \sum_{k=1}^{n} (a_k+b_k) &= \lim_{n\to\infty} \left(\sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k\right) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} a_k + \lim_{n\to\infty} \sum_{k=1}^{n} b_k \\
&= a+b\,.
\end{aligned}$$

2.4 Ejemplo. Demuéstrese que $\sum_{k=1}^{\infty} \frac{3}{2^k} = 3\,.$

Solución. Como $\sum_{k=1}^{\infty} \frac{3}{2^k} = \sum_{k=0}^{\infty} \frac{3}{2}\cdot\frac{1}{2^k}$ y $\sum_{k=0}^{\infty} \frac{1}{2^k}$ tiene la suma 2,

$$\sum_{k=0}^{\infty} \frac{3}{2}\cdot\frac{1}{2^k} = \frac{3}{2}\cdot 2 = 3\,.$$

Problemas

1. Determínese si las siguientes series convergen o divergen. Proporciónense las sumas de las que converjan.

a) $\sum_{k=0}^{\infty} (-1)^k \frac{1}{3^k}$ *b*) $\sum_{k=1}^{\infty} \frac{4}{3^k}$

c) $\sum_{k=0}^{\infty} \frac{4^k}{3^k}$ *d*) $\sum_{k=2}^{\infty} \frac{2}{5^k}$

e) $\sum_{k=0}^{\infty} \frac{1}{2^{2k+1}}$ *f*) $\sum_{k=0}^{\infty} \frac{2^k+3^k}{5^k}$.

2. Determínese si las siguientes series de números complejos convergen o divergen. Proporciónese la suma de las que converjan.

a) $\sum_{k=1}^{\infty} \left(\frac{1}{2^k} + i2^k\right)$ *b*) $\sum_{k=0}^{\infty} \left(\frac{1}{4^k} + \frac{2}{3^k} i\right)$.

3. Demuéstrese que $\sum_{k=1}^{\infty} \frac{1}{k^2+k} = 1$.

Sugerencia: $s_n = \sum_{k=1}^{n} \frac{1}{k^2+k} = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right)$

4. Demuéstrese que $\sum_{k=1}^{\infty} \frac{k-1}{2^{k+1}} = \frac{1}{2}$.

Sugerencia: $s_n = \sum_{k=1}^{n} \frac{k-1}{2^{k+1}} = \sum_{k=1}^{n} \left(\frac{k}{2^k} - \frac{k+1}{2^{k+1}}\right)$.

5. Pruébese que $\sum_{k=1}^{\infty} \ln \frac{k}{k+1}$ diverge.

6. Pruébese que si Σa_k diverge y $c \neq 0$, entonces $\Sigma c a_k$ diverge.

7. Pruébese que si Σa_k diverge y Σb_k converge, entonces $\Sigma(a_k + b_k)$ diverge.

8. Proporciónense ejemplos donde Σa_k y Σb_k divergen y

a) $\Sigma(a_k + b_k)$ converge; *b*) $\Sigma(a_k + b_k)$ diverge.

9. ¿Converge $\sum_{k=0}^{\infty} \frac{5^k + 3^k}{4^k}$?

10. ¿Converge $\sum_{k=1}^{\infty} (\sqrt{k+1} - \sqrt{k})$?

3. PRUEBAS DE CONVERGENCIA Y DIVERGENCIA DE SERIES

De acuerdo con las definiciones dadas, si queremos saber si la serie $\sum_{k=1}^{\infty} a_k$ converge, todo lo que tenemos que hacer es investigar la convergencia de la sucesión $\{s_n\}$ donde $s_n = \sum_{k=1}^{n} a_k$. Sin embargo, en muchos casos no somos capaces de obtener una expresión para s_n que nos sirva para poder determinar la convergencia de $\{s_n\}$. Es, pues, conveniente desarrollar criterios para la convergencia de la serie Σa_k en términos de la sucesión $\{a_k\}$. Es esto lo que haremos.

Obtenemos primero una prueba de divergencia. Si la serie Σa_k converge a a, entonces

$$\text{lím } a_k = \text{lím } (s_k - s_{k-1}) = a - a = 0.$$

Así pues, si Σa_k converge, entonces lím $a_k = 0$. Este hecho da lugar a la siguiente prueba de divergencia.

3.1 *Si $\{a_k\}$ no converge a cero, entonces Σa_k diverge.*

3.2 Ejemplo. ¿Converge $\Sigma \dfrac{k^2+3k}{2k^2+5}$?

SOLUCIÓN. Como lím $\dfrac{k^2+3k}{2k^2+5} = \dfrac{1}{2}$, la serie diverge.

Debe tenerse cuidado en no creer que Σa_k converge si lím $a_k = 0$. Esto no es necesariamente cierto. Por ejemplo, $\Sigma \dfrac{1}{k}$ diverge (lo que probaremos más tarde), pero lím $\dfrac{1}{k} = 0$.

Si los términos de una serie Σa_k son no negativos, entonces la sucesión $\{s_n\}$, donde $s_n = \sum_{k=1}^{n} a_k$, es una sucesión no decreciente. De donde, la sucesión $\{s_n\}$ (es decir, la serie Σa_k) o converge o diverge a ∞ según que esté acotada o no. Por tanto, si $a_k \geqslant 0$, los criterios para conocer si la sucesión $\{s_n\}$ es o no acotada, constituyen un criterio para la convergencia de la serie Σa_k. La prueba básica para la convergencia y divergencia de series con términos no negativos es el *criterio de comparación* dado en el teorema 3.3 y en el corolario 3.4 que a continuación presentamos.

3.3 Teorema. *Si Σa_k y Σb_k son series con términos no negativos, si Σb_k converge y si $a_k \leqslant b_k$ para todo k suficientemente grande, entonces Σa_k converge.*

PRUEBA. Supongamos $a_k \leqslant b_k$ para todo k mayor que cierto entero positivo N y $\Sigma b_k = b$. Entonces, para todo $n > N$,

$$s_n = \sum_{k=1}^{n} a_k = \sum_{k=1}^{N} a_k + \sum_{k=N+1}^{n} a_k \leqslant \sum_{k=1}^{N} a_k + \sum_{k=N+1}^{n} b_k \leqslant \sum_{k=1}^{N} a_k + b\,.$$

Como $b \geqslant 0$, $s_n \leqslant \sum_{k=1}^{N} a_k + b$ para todo n y, por tanto, $\{s_n\}$ es acotada. Luego Σa_k converge.

3.4 Corolario. *Si Σa_k y Σb_k son series con términos no negativos, si Σb_k diverge y si $a_k \geqslant b_k$ para todo k suficientemente grande, entonces Σa_k diverge.*

PRUEBA. Supongamos que Σa_k converge. Entonces, según el teorema 3.3, Σb_k converge en contra de la hipótesis. Lo que prueba el corolario.

Otra forma del *criterio de comparación* que es generalmente más fácil de aplicar que la de 3.3 o 3.4 es:

3.5 Corolario. *Supongamos que Σa_k es una serie de términos no negativos y Σb_k es una serie de términos positivos.*

1) *Si* $\lim \dfrac{a_k}{b_k} = c \geqslant 0$ *y Σb_k converge, entonces Σa_k converge.*

2) *Si* $\lim \dfrac{a_k}{b_k} = c > 0$ *o* ∞ *y Σb_k diverge, entonces Σa_k diverge.*

PRUEBA. Solamente probaremos 1); la prueba de 2) es del todo análoga. Como $\lim \dfrac{a_k}{b_k} = c$, existe un número N tal que $\dfrac{a_k}{b_k} < c+1$ siempre que $k > N$.

Si Σb_k converge, entonces $\Sigma(c+1)b_k$ converge y, por tanto, Σa_k converge, ya que $a_k < (c+1)b_k$ para todo k suficientemente grande.

3.6 Ejemplo. Determínese si la serie $\Sigma \dfrac{k^3}{2^k}$ converge.

SOLUCIÓN. Determinamos primero el límite de $\{a_k\}$:

$$\lim \frac{k^3}{2^k} = 0\,.$$

Aquí pues, 3.1 no se aplica. Para k grande, $\left(\dfrac{3}{2}\right)^k$ es mayor que k^3. Por tanto, para k suficientemente grande, $\dfrac{k^3}{2^k}$ es menor que $\left(\dfrac{3}{4}\right)^k$. Comparando $\Sigma \dfrac{k^3}{2^k}$ con

la serie $\Sigma\left(\frac{3}{4}\right)^k$ que converge, tenemos

$$\lim \frac{k^3}{2^k} \Big/ \left(\frac{3^k}{4}\right) = \lim \frac{k^3}{(\frac{3}{2})^k} = 0.$$

Por tanto, $\Sigma \frac{k^3}{2^k}$ converge de acuerdo con el criterio de comparación (corolario 3.5).

Aunque el criterio de comparación es un criterio de convergencia para series con términos no negativos, podemos a veces usarlo para probar la convergencia de otras series. Si Σa_k es una serie cualquiera de números reales, entonces $\Sigma |a_k|$ es una serie de términos no negativos y, por tanto, el criterio de comparación puede aplicarse a $\Sigma |a_k|$. A continuación probaremos que si $\Sigma |a_k|$ converge, entonces Σa_k converge también.

3.7 Teorema. *Si Σa_k converge, entonces $\Sigma |a_k|$ converge.*

PRUEBA. Supongamos que $\Sigma |a_k|$ converge. Como $-|a_k| \leqslant a_k \leqslant |a_k|$,

$$0 \leqslant a_k + |a_k| \leqslant 2|a_k|.$$

Así pues, $\Sigma(a_k + |a_k|)$ es una serie con términos no negativos cada uno de cuyos términos es menor que o igual al término correspondiente de la serie convergente $\Sigma 2|a_k|$. Por tanto, según el criterio de comparación (teorema 3.3) $\Sigma(a_k + |a_k|)$ converge. Luego $\Sigma a_k = \Sigma(a_k + |a_k| - |a_k|)$ converge de acuerdo con el teorema 2.3. Esto completa la prueba.

Decimos que la serie Σa_k es *absolutamente convergente* si $\Sigma |a_k|$ converge. Así pues, el teorema 3.7 nos dice que una serie absolutamente convergente es convergente. Es posible, como más tarde veremos, que Σa_k converja aunque $\Sigma |a_k|$ diverja. Si Σa_k converge, pero $\Sigma |a_k|$ diverge, entonces decimos que Σa_k es *condicionalmente convergente.*

La utilidad del criterio de comparación para determinar la convergencia de una serie dada depende de que tengamos cierta cantidad de series convergentes o divergentes conocidas con las que podamos comparar las series dadas. Hasta el momento tenemos esencialmente sólo un tipo de serie convergente o divergente conocida: la serie geométrica. Sin embargo, comparando series con series de este tipo (geométricas) podemos desarrollar algunos criterios de convergencia muy útiles.

3.8 Teorema. (Criterio de la razón.) *Si $a_k \neq 0$*

1) $\overline{\lim} \left|\frac{a_{k+1}}{a_k}\right| < 1$ *implica que Σa_k es absolutamente convergente,*

2) $\underline{\lim}\left|\dfrac{a_{k+1}}{a_k}\right| > 1$ *implica que* Σa_k *es divergente.*

PRUEBA

1) Tomemos una r tal que $\overline{\lim}\left|\dfrac{a_{k+1}}{a_k}\right| < r < 1$. Entonces existe un número N tal que $\left|\dfrac{a_{k+1}}{a_k}\right| \leqslant r$ siempre que $k > N$. Es decir,

$$\left|\frac{a_{k+1}}{a_k}\right| \leqslant \frac{r^{k+1}}{r^k}$$

o lo que es equivalente,

$$\frac{|a_{k+1}|}{r^{k+1}} \leqslant \frac{|a_k|}{r_k} \quad \text{siempre que } k > N.$$

Por tanto, para k suficientemente grande, $\left\{\dfrac{|a_k|}{r^k}\right\}$ es una sucesión no creciente de términos positivos y, por tanto, es acotada. Así pues, existe un número b tal que, para toda k, $\dfrac{|a_k|}{r^k} \leqslant b$, o lo que es lo mismo,

$$|a_k| \leqslant br^k.$$

Como $|r| < 1$, Σbr^k converge y, por tanto, $\Sigma |a_k|$ converge según el criterio de comparación. Y esto prueba la parte 1).

2) Si $\underline{\lim}\left|\dfrac{a_{k+1}}{a_k}\right| > 1$, entonces existe un número N tal que $\left|\dfrac{a_{k+1}}{a_k}\right| > 1$ siempre que $k > N$; es decir,

$$|a_{k+1}| > |a_k| \quad \text{siempre que } k > N.$$

Así pues, para k suficientemente grande, $\{|a_k|\}$ es una sucesión creciente de términos positivos. Por tanto, $\{a_k\}$ no puede convergir a cero y, por tanto, Σa_k diverge según 3.1, pág. 497.

Si $\left\{\left|\dfrac{a_{k+1}}{a_k}\right|\right\}$ tiene un límite, entonces podemos enunciar el criterio de la razón en la siguiente forma.

3.9 Corolario. *Supongamos que* $a_k \neq 0$ *y que* $\lim\left|\dfrac{a_{k+1}}{a_k}\right| = L$. *Si* $L < 1$, *entonces* Σa_k *es absolutamente convergente. Si* $L > 1$, *entonces* Σa_k *es divergente.*

3.10 Ejemplo. Pruébese que $\Sigma \frac{k^3}{k!}$ converge.

SOLUCIÓN.

$$\lim \left[\frac{(k+1)^3}{(k+1)!} \Big/ \frac{k^3}{k!} \right] = \lim \left[\frac{(k+1)^3}{(k+1)!} \cdot \frac{k!}{k^3} \right] = \lim \left[\left(\frac{k+1}{k} \right)^2 \cdot \frac{1}{k} \right] = 0 .$$

Por tanto, $\Sigma \frac{k^3}{k!}$ converge por el criterio de la razón.

3.11 Teorema. (Criterio de la raíz.)

1) *Si* $\overline{\lim}\ \sqrt[k]{|a_k|} < 1$, *entonces* Σa_k *es absolutamente convergente.*

2) *Si* $\overline{\lim}\ \sqrt[k]{|a_k|} > 1$, *entonces* Σa_k *es divergente.*

PRUEBA.

1) Tomemos un número r tal que $\overline{\lim}\ \sqrt[k]{|a_k|} < r < 1$. Entonces existe un número N tal que $\sqrt[k]{|a_k|} \leqslant r$ siempre que $k > N$. Es decir

$$|a_k| \leqslant r^k \quad \text{siempre que } k > N.$$

Como $r < 1$, Σr^k converge y, por tanto, $\Sigma |a_k|$ converge según el criterio de comparación.

2) Si $\overline{\lim}\ \sqrt[k]{|a_k|} > 1$, entonces $\sqrt[k]{|a^k|} > 1$ para infinitos k. Es decir, $|a_k| > 1$ para infinitos k y, $\{a_k\}$ no converge a cero. Y esto prueba que Σa_k diverge.

Si $\{\sqrt[k]{|a_k|}\}$ tiene un límite, entonces el criterio de la raíz puede enunciarse en la siguiente forma.

3.12 Corolario. *Supongamos que* $\lim \sqrt[k]{|a_k|} = L$. *Si* $L < 1$ *entonces* Σa_k *es absolutamente convergente. Si* $L > 1$, *entonces* Σa_k *es divergente.*

3.13 Ejemplo. ¿Converge la serie $\sum_{k=1}^{\infty} \frac{k^4}{2^k}$?

SOLUCIÓN. Aplicando el criterio de la raíz tenemos

$$\lim \sqrt[k]{\frac{k^4}{2^k}} = \lim \frac{(\sqrt[k]{k})^4}{2} = \frac{1}{2}. \quad \text{(ejemplo 3.12, pág. 461.)}$$

Por tanto, $\Sigma \frac{k^4}{2^k}$ converge.

Si al aplicar el criterio de la razón, $\lim \left| \frac{a_{k+1}}{a_k} \right| = 1$, o al aplicar el criterio

de la raíz, $\lim \sqrt[k]{|a_k|} = 1$, entonces estos criterios no nos dan ninguna información respecto a la convergencia de Σa_k. Por ejemplo, consideremos la serie $\Sigma \frac{1}{k^r}$ donde r es cierto número real. Como se probará en el próximo ejemplo 3.15, si $r > 1$ esta serie converge mientras que si $r \leqslant 1$ la serie diverge. Sin embargo,

$$\lim \left| \frac{a_{k+1}}{a_k} \right| = \lim \left(\frac{k}{k+1} \right)^r = 1$$

y

$$\lim \sqrt[k]{|a_k|} = \lim \left(\frac{1}{\sqrt[k]{k}} \right)^r = 1 .$$

Por tanto, la convergencia de $\Sigma \frac{1}{k^r}$ no puede determinarse mediante la aplicación del criterio de la razón o el criterio de la raíz.

Hemos estado enunciando criterios de convergencia para series de números reales ya que una serie de puntos en R^m converge si y sólo si las series componentes convergen. Sin embargo, es conveniente señalar que los criterios de comparación, razón y raíz pueden aplicarse directamente a series de puntos. Por ejemplo, tenemos el siguiente criterio de comparación:

Si $\Sigma \mathbf{a}_k$ *es una serie de puntos en* R^m *y* Σb_k *es una serie de números reales que converge si* $|\mathbf{a}_k| \leqslant b_k$ *para todo* k *suficientemente grande, entonces* $\Sigma \mathbf{a}_k$ *converge*. Puede verse fácilmente esto como sigue. Si $\mathbf{a}_k = (a_k{}^1, \ldots, a_k{}^m)$ entonces $|a_k{}^j| \leqslant |\mathbf{a}_k|$ $(j = 1, \ldots, m)$. Así pues $|a_k{}^j| \leqslant |b_k|$ para todo k suficientemente grande y, por tanto, cada una de las series componentes $\Sigma a_k{}^j$ converge según el teorema 3.3. Luego la serie de puntos $\Sigma \mathbf{a}_k$ converge.

Las otras formas del criterio de comparación y de los criterio de la razón y la raíz para series de puntos se deducen fácilmente del anterior criterio de comparación. Todo lo que se necesita es reemplazar el valor absoluto $|\mathbf{a}_k|$

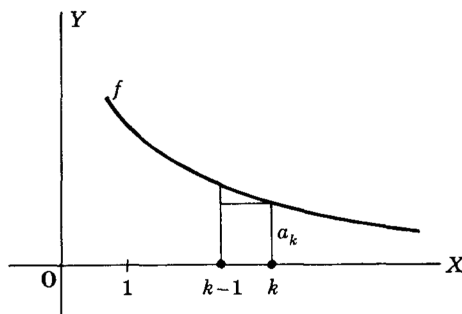


FIGURA 1

por la longitud del vector $|\mathbf{a}_k|$ en el enunciado y prueba de cada uno de estos teoremas.

Damos ahora otro criterio de convergencia para series de números reales en el que comparamos una serie con una integral impropia relacionada con ella. Podremos determinar la convergencia de $\Sigma \frac{1}{k^r}$ aplicando este criterio.

3.14 Teorema. (Criterio de la integral.) *Si Σa_k es una serie de términos no negativos y f es una función continua no creciente sobre el intervalo* $[1, \infty\rangle$ *tal que $f(k) = a_k$, entonces Σa_k e $\int_1^\infty f$ o ambas, convergen o ambas, divergen.*

PRUEBA. Recuérdese que la integral impropia $\int_1^\infty f$ está definida como el $\lim\limits_{b\to\infty} \int_1^b f$. Así pues, si definimos $g(b) = \int_1^b f$, entonces $\int_1^\infty f = \lim\limits_{b\to\infty} g(b)$. Como sobre $[1, \infty\rangle$ los valores de f son no negativos, g es una función no decreciente. Luego, $\lim\limits_{b\to\infty} g(b) = c$ o $\lim\limits_{b\to\infty} g(b) = \infty$.

1) Supongamos que $\int_1^\infty f$ converge a c. Como f es una función no creciente (figura 1), si $k \geqslant 2$ entonces $a_k = f(k) \leqslant f(x)$ para $x \in [k-1, k]$ y, por tanto,

$$a_k = \int_{k-1}^k a_k \leqslant \int_{k-1}^k f.$$

Entonces

$$s_n = \sum_{k=1}^n a_k \leqslant a_1 + \sum_{k=2}^n \int_{k-1}^k f = a_1 + \int_1^n f \leqslant a_1 + c.$$

Así pues, $\{s_n\}$ es una sucesión no decreciente acotada y, por tanto, Σa_k converge.

2) Supongamos que $\int_1^\infty f$ diverge a ∞. Como

$$a_k = f(k) \geqslant f(x) \quad \text{para} \quad x \in [k, k+1],$$

$$a_k = \int_k^{k+1} a_k \geqslant \int_k^{k+1} f.$$

Entonces

$$s_n = \sum_{k=1}^n a_k \geqslant \sum_{k=1}^n \int_k^{k+1} f = \int_1^{n+1} f$$

y, como $\int_1^\infty f$ diverge a ∞, la sucesión $\{s_n\}$ diverge a ∞. Por tanto, Σa_k diverge. Y esto completa la prueba.

3.15 Ejemplo. Demuéstrese que $\sum_{k=1}^{\infty} \frac{1}{k^r}$ converge si $r > 1$ y diverge si $r \leqslant 1$.

SOLUCIÓN. Si $r \leqslant 0$, entonces $\left\{\frac{1}{k^r}\right\}$ no converge a cero y, por tanto, $\Sigma \frac{1}{k^r}$ diverge según 3.1. Si $r > 0$, sea $f(x) = \frac{1}{x^r}$. La función f es una función continua decreciente sobre $[1, \infty\rangle$ tal que $f(k) = \frac{1}{k^r}$. Por tanto, si $r > 0$, $\Sigma \frac{1}{k^r}$ converge si y sólo si $\int_1^{\infty} \frac{dx}{x^r}$ converge. Como

$$\int_1^b \frac{dx}{x^r} = \begin{cases} \frac{1}{r-1}(1 - b^{1-r}) & \text{si } r \neq 1 \\ \ln b & \text{si } r = 1, \end{cases}$$

$\int_1^{\infty} \frac{dx}{x^r}$ converge si $r > 1$ y diverge si $r \leqslant 1$. Así pues, $\Sigma \frac{1}{k^r}$ converge si $r > 1$ y diverge si $r \leqslant 1$.

En particular, $\Sigma \frac{1}{k}$ diverge y $\Sigma \frac{1}{k^2}$ converge. Tanto en una como en otra de estas series el término general tiende a cero; en el caso de $\Sigma \frac{1}{k^2}$ el término se hace pequeño con suficiente rapidez para que las sumas s_n estén acotadas, en el caso de $\Sigma \frac{1}{k}$, la rapidez no es suficiente para que se produzca la acotación de las sumas parciales.

Las series del tipo $\Sigma \frac{1}{k^r}$ junto con las series geométricas Σx^k proveen una reserva adecuada de series de carácter convergente o divergente conocido para su uso en el criterio de comparación. Hay muchos más criterios para la convergencia de series de términos no negativos que los que aquí podemos dar, pero los que discutamos serán suficientes para nuestros propósitos.[1] Como señalamos antes, los criterios para la convergencia de series con términos no negativos son también criterios para la convergencia absoluta de series con términos de signo arbitrario. Por ejemplo, $\Sigma(-1)^k \frac{1}{k^2}$ es absolutamente convergente, ya que $\Sigma\left|(-1)^k \frac{1}{k^2}\right| = \Sigma \frac{1}{k^2}$ converge.

[1] Una discusión extensiva puede encontrarse en la referencia [40].

Consideremos ahora la serie $\Sigma(-1)^{k+1}\frac{1}{k}$. Como $\Sigma\frac{1}{k}$ diverge, $\Sigma(-1)^{k+1}\frac{1}{k}$ no es absolutamente convergente. Sin embargo, la serie $\Sigma(-1)^{k+1}\frac{1}{k}$ puede convergir; es decir, puede ser condicionalmente convergente. Calculando algunas de las sumas s_n, tenemos:

$$s_1 = 1,\ s_2 = 1-\tfrac{1}{2} = \tfrac{1}{2},\ s_3 = s_2+\tfrac{1}{3} = \tfrac{5}{6},\ s_4 = s_3-\tfrac{1}{4} = \tfrac{7}{12},$$

$$s_5 = s_4+\tfrac{1}{5} = \tfrac{47}{60},\ s_6 = s_5-\tfrac{1}{6} = \tfrac{37}{60}.$$

Si trazamos estas sumas sobre una recta (figura 2), la forma en que aparecen distribuidas sugiere que la serie converge: las sumas avanzan y retroceden en la recta y cada vez se mueven una distancia más corta.

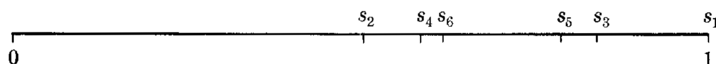


FIGURA 2

Mirándolo desde otro punto de vista, parece que las sumas pares forman una sucesión acotada superiormente por s_1 y las sumas impares una sucesión decreciente inferiormente acotada por s_2. De donde cada una de estas subsucesiones de $\{s_n\}$ converge y, como los términos de estas subsucesiones se aproximan cada vez más, ambas deben convergir al mismo punto. Podríamos probar todas estas afirmaciones, pero en lugar de hacerlo probamos un teorema general que demuestra la convergencia de esta serie como un caso particular. La prueba del teorema sigue los lineamientos que acabamos de exponer.

3.16 Teorema. (Criterio de las series alternantes.) *Si $\{a_k\}$ es una sucesión no creciente de términos positivos y* lím $a_k = 0$, *entonces* $\sum_{k=1}^{\infty}(-1)^{k+1}a_k$ *converge.*

PRUEBA. Como

$$s_{2n+2} = s_{2n}+a_{2n+1}-a_{2n+2} \geqslant s_{2n}$$

y

$$s_{2n+1} = s_{2n-1}-a_{2n}+a_{2n+1} \leqslant s_{2n-1},$$

$\{s_{2n}\}$ es una sucesión no decreciente y $\{s_{2n-1}\}$ es una sucesión no creciente. Además

$$s_2 \leqslant s_{2n} = s_{2n-1}-a_{2n} < s_{2n-1} \leqslant s_1.$$

Así pues, $\{s_{2n}\}$ está acotada superiormente por s_1 y $\{s_{2n-1}\}$ está acotada inferiormente por s_2. Por tanto, $\{s_{2n}\}$ y $\{s_{2n-1}\}$ convergen. Como

$$\text{lím } s_{2n-1} - \text{lím } s_{2n} = \text{lím }(s_{2n-1} - s_{2n}) = \text{lím } a_{2n} = 0,$$

estas dos sucesiones convergen al mismo punto, llamémosle c. Luego, lím $s_n = c$ (problema 10, pág. 457). Y esto completa la prueba.

3.17 Ejemplo. ¿Converge la serie $\sum_{k=1}^{\infty} (-1)^{k+1} \dfrac{k-2}{k^2+3k}$?

SOLUCIÓN. Si hacemos $a_k = \dfrac{k-2}{k^2+3k}$, tenemos lím $a_k = 0$. Si $\{a_k\}$ es una sucesión no creciente de términos positivos, entonces el criterio de las series alternantes demuestra que la serie converge. Claramente, los términos a_k son positivos para $k > 2$. Además,

$$\begin{aligned} a_{k+1} \leqslant a_k &\Leftrightarrow \frac{k-1}{k^2+5k+4} \leqslant \frac{k-2}{k^2+3k} \\ &\Leftrightarrow k^3+2k^2-3k \leqslant k^3+3k^2-6k-8 \\ &\Leftrightarrow k^2-3k-8 \geqslant 0 \\ &\Leftrightarrow k \geqslant 5. \end{aligned}$$

Por tanto, $\{a_k\}$ es una sucesión no creciente de términos positivos para $k \geqslant 5$. Como la convergencia de una sucesión y, por tanto, de una serie, depende solamente del comportamiento de los términos desde algún punto en adelante, $\Sigma(-1)^{k+1} \dfrac{k-2}{k^2+3k}$ converge de acuerdo con el criterio de las series alternantes.

Problemas

1. Determínese si las siguientes series convergen o divergen.

a) $\sum_{k=1}^{\infty} \dfrac{k^2-3}{3k^2+4k}$ *b*) $\sum_{k=1}^{\infty} \dfrac{2^k}{3^k+k}$

c) $\sum_{k=1}^{\infty} \dfrac{3^k-2}{2^k+5k}$ *d*) $\sum_{k=1}^{\infty} (-1)^k \dfrac{k^2+4}{2k^2+3k}$

e) $\sum_{k=1}^{\infty} \dfrac{k^2}{3^k}$ *f*) $\sum_{k=1}^{\infty} \dfrac{k^3+3k}{5^k+2}$.

2. Si $a_k \geqslant 0$ y $\lim_{k\to\infty} r^k a_k = 0$ para algún $r > 1$, pruébese que Σa_k converge.

3. Si Σa_k es absolutamente convergente, pruébese que Σa_k^2 es convergente.

4. Determínese si las siguientes series convergen o divergen.

a) $\sum_{k=1}^{\infty} \frac{2^k}{k!}$ *b)* $\sum_{k=1}^{\infty} \frac{k!}{k^4+3}$

c) $\sum_{k=1}^{\infty} \frac{3^k+4k^2}{k!+7k}$ *d)* $\sum_{k=1}^{\infty} \frac{k!}{k^k}$

e) $\sum_{k=1}^{\infty} \frac{(k+1)^k}{e^{k^2}}$ *f)* $\sum_{k=1}^{\infty} (-1)^k \frac{k^3+5k}{3^k+4}$.

5. Pruébese que $\sum_{k=2}^{\infty} \frac{1}{k(\ln k)^r}$ converge cuando $r > 1$ y diverge cuando $r \leqslant 1$.

6. Determínese si las siguientes series convergen o divergen.

a) $\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}$ *b)* $\sum_{k=1}^{\infty} \frac{\ln k}{k}$

c) $\sum_{k=2}^{\infty} \frac{1}{\sqrt{k}\ln k}$ *d)* $\sum_{k=1}^{\infty} \frac{\sqrt{k}}{3k^2-1}$

e) $\sum_{k=4}^{\infty} \frac{1}{k \ln k \ln \ln k}$ *f)* $\sum_{k=2}^{\infty} \frac{1}{(\ln k)^k}$

g) $\sum_{k=1}^{\infty} \frac{k+7}{3k^2+k-2}$ *h)* $\sum_{k=1}^{\infty} \frac{k^3-5k}{4k^4+7k^3+9}$

i) $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{k+2}{3k^2+5}$ *j)* $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{k^2+4}{2k^3+3k}$

k) $\sum_{k=1}^{\infty} (-1)^k \frac{k-3}{4k+7}$ *l)* $\sum_{k=1}^{\infty} \frac{\cos k\pi}{k^2}$

m) $\sum_{k=1}^{\infty} (-1)^k \frac{\ln k}{k}$ *n)* $\sum_{k=1}^{\infty} \frac{\ln k}{k^2}$.

7. Si $a_k \neq 0$ y $\left|\frac{a_{k+1}}{a_k}\right| \leqslant 1 - \frac{2}{k} + \frac{1}{k^2}$, pruébese que Σa_k es absolutamente convergente.

8. Si Σa_k^2 converge, pruébese que $\Sigma \frac{a_k}{k}$ converge.

9. En la prueba del teorema 3.8 probamos que si $\overline{\lim}\left|\frac{a_{k+1}}{a_k}\right| < 1$, entonces para cualquier número r tal que $\overline{\lim}\left|\frac{a_{k+1}}{a_k}\right| < r < 1$ existe un número b tal que $|a_k| \leqslant br^k$ para todo k. Conclúyase de esto que para cualquier número s tal que $\overline{\lim}\left|\frac{a_{k+1}}{a_k}\right| < s < 1$, $|a_k| \leqslant s^k$ para toda k suficientemente grande.

4. LA SUMA DE UNA SERIE CONVERGENTE

Cuando probamos que la serie geométrica $\sum_{k=0}^{\infty} x^k$ converge para $|x| < 1$, obtuvimos la suma de la serie: $\frac{1}{1-x}$. Sin embargo, si usamos uno de los criterios discutidos en la sección precedente para mostrar la convergencia de una serie, no obtenemos la suma; sabemos que la sucesión $\{s_n\}$ tiene un límite, pero no sabemos cuál es éste. Sabemos, sin embargo, que podemos aproximarnos a la suma tanto como deseemos con las s_n tomando n suficientemente grande. En esta sección daremos algunas estimaciones de cuán grande debe ser n para que s_n se aproxime a la suma de la serie con un grado especificado de precisión.

Si probamos que la serie Σa_k es absolutamente convergente comparándola con la serie convergente Σb_k, entonces podemos obtener una estimación del error que se comete al usar s_n como una aproximación a la suma a, como sigue. Si $|a_k| \leqslant b_k$ siempre que $k > N$, entonces $m > n \geqslant N$ implica

$$|s_m - s_n| = \left|\sum_{k=n+1}^{m} a_k\right| \leqslant \sum_{k=n+1}^{m} |a_k| \leqslant \sum_{k=n+1}^{m} b_k .$$

Por tanto,

4.1 $$|a - s_n| = \lim_{m\to\infty} |s_m - s_n| \leqslant \sum_{k=n+1}^{\infty} b_k \quad \text{siempre que } n \geqslant N.$$

Al número $|a - s_n|$ se le llama *error de truncación.*

Recuérdese que, al probar la convergencia de una serie mediante el criterio de la razón o el de la raíz, estamos en realidad comparando la serie dada con una serie geométrica. Si $\overline{\lim}\left|\frac{a_{k+1}}{a_k}\right| = c < 1$ o $\overline{\lim}\sqrt[k]{|a_k|} = c < 1$, entonces, tomando cualquier número r tal que $c < r < 1$, tenemos $|a_k| \leqslant r^k$ para todo k suficientemente grande.

4.2 Ejemplo. Demuéstrese que $\sum_{k=1}^{\infty} \frac{k^2}{3^k}$ converge y estímese a partir de qué n las sumas s_n se aproximan al valor de la serie con un error de truncación menor que 1×10^{-4}.

SOLUCIÓN. Usando el criterio de la razón, tenemos

$$\text{lím} \frac{(k+1)^2}{3^{k+1}} \bigg/ \frac{k^2}{3^k} = \text{lím} \frac{1}{3}\left(\frac{k+1}{k}\right)^2 = \frac{1}{3}.$$

Así pues, $\Sigma \frac{k^2}{3^k}$ converge y $\frac{k^2}{3^k}$ será menor que $\frac{1}{2^k}$ para k suficientemente grande. En realidad,

$$\frac{k^2}{3^k} < \frac{1}{2^k} \quad \text{siempre que } k \geqslant 13.$$

Si llamamos a la suma de la serie a, entonces el error de truncación que resulta de aproximar a a por s_n es $|a - s_n|$. De acuerdo con 4.1

$$|a - s_n| \leqslant \sum_{k=n+1}^{\infty} \frac{1}{2^k} = \frac{1}{2^{n+1}} \sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{2^n} \quad \text{siempre que } n \geqslant 13.$$

Si tomamos $n = 14$, entonces $\frac{1}{2^n} < 1 \times 10^{-4}$. Así pues, $s_{14} = \sum_{k=1}^{14} \frac{k^2}{3^k}$ se aproxima a la suma de la serie con un error menor que 1×10^{-4}.

Supongamos que hemos probado que Σa_k es absolutamente convergente según el criterio de la integral. Si f es una función continua no creciente sobre $[1, \infty\rangle$ tal que $f(k) = |a_k|$, entonces para $m > n$

$$|s_m - s_n| = \left| \sum_{k=n+1}^{m} a_k \right| \leqslant \sum_{k=n+1}^{m} |a_k| \leqslant \sum_{k=n+1}^{m} \int_{k-1}^{k} f = \int_{n}^{m} f.$$

Por tanto, si la suma de la serie es a, obtenemos la siguiente cota para el error de truncación:

4.3 $$|a - s_n| \leqslant \int_{n}^{\infty} f.$$

4.4 Ejemplo. Hállese el valor aproximado de la suma de $\sum_{k=1}^{\infty} \frac{1}{k^2}$ con un error de truncación que no exceda a 0.1.

SOLUCIÓN. $\Sigma \frac{1}{k^2}$ converge, ya que $\int_{1}^{\infty} \frac{dx}{x^2}$ converge. Si la suma de la serie

es a, entonces de acuerdo con 4.3

$$|a-s_n| \leqslant \int_n^{\infty} \frac{dx}{x^2} = \frac{1}{n}.$$

Luego, $|a-s_n|$ no excederá a 0.1 si $n = 10$. Calculando $s_{10} = \sum_{k=1}^{10} \frac{1}{k^2}$, tenemos

$$1.54 < s_{10} < 1.55.$$

Como $\Sigma \frac{1}{k^2}$ es una serie con términos positivos, a es mayor que s_{10} y, por tanto,

$$1.54 < a < 1.65.$$

De donde a es aproximadamente 1.6.

Si Σa_k es una serie de términos no negativos que se conoce es convergente mediante el criterio de la integral, entonces nos encontramos con la acotación más estricta

$$\int_{n+1}^{\infty} f \leqslant a-s_n \leqslant \int_n^{\infty} f.$$

Así pues, en el ejemplo 4.4 tenemos

$$\frac{1}{n+1} \leqslant a-s_n \leqslant \frac{1}{n}$$

y, para $n = 10$,

$$1.63 < 1.54 + \tfrac{1}{11} < a < 1.55 + \tfrac{1}{10} = 1.65.$$

Si $\{a_k\}$ es una sucesión no creciente de términos positivos y $\text{lím}\, a_k = 0$, entonces $\sum_{k=1}^{\infty} (-1)^{k+1} a_k$ vemos que converge de acuerdo con el criterio de la serie alternante. Además, si a es la suma de la serie, sabemos que $a = \sup\{s_{2n}\} = \text{ínf}\{s_{2n-1}\}$. Luego, para cualquier n,

$$s_{2n} < a < s_{2n+1} \quad \text{y} \quad s_{2n} < a < s_{2n-1};$$

es decir

$$0 < a-s_{2n} < s_{2n+1}-s_{2n} = a_{2n+1}$$

y

$$0 > a-s_{2n-1} > s_{2n}-s_{2n-1} = -a_{2n}.$$

Luego, para cualquier n,

4.5 $$|a-s_n| < a_{n+1}.$$

Es decir, si aproximamos la suma a mediante una suma finita s_n, entonces el error de truncación es menor que el valor absoluto del primer término que no se ha tenido en cuenta.

4.6 Ejemplo. Pruébese que $\Sigma(-1)^{k+1}\dfrac{k-2}{3k^2+4k}$ converge y estímese a partir de qué valor de n la suma parcial s_n se aproxima a la suma con un error de truncación menor que 0.01.

SOLUCIÓN. Sea $a_k = \dfrac{k-2}{3k^2+4k}$. Entonces $a_k > 0$ si $k \geqslant 3$ y $\lim a_k = 0$. Ademas,

$$\begin{aligned} a_{k+1} \leqslant a_k &\Leftrightarrow \frac{k-1}{3k^2+10k+7} \leqslant \frac{k-2}{3k^2+4k} \\ &\Leftrightarrow 3k^2-9k-14 \geqslant 0 \\ &\Leftrightarrow k \geqslant 5\,. \end{aligned}$$

Así pues, para $k \geqslant 5$, $\{a_k\}$ es una sucesión no creciente de términos positivos que converge a cero. Luego $\Sigma(-1)^{k+1}\dfrac{k-2}{3k^2+4k}$ converge de acuerdo con el criterio de la serie alternante. Si la suma de la serie es a, entonces

$$|a-s_n| < a_{n+1} = \frac{n-1}{3n^2+10n+7} \quad \text{para} \quad n \geqslant 5\,.$$

Tenemos

$$\begin{aligned} \frac{n-1}{3n^2+10n+7} < \frac{1}{100} &\Leftrightarrow 3n^2-90n+107 > 0 \\ &\Leftrightarrow 3(n-15)^2-568 > 0 \\ &\Leftrightarrow n \geqslant 29\,. \end{aligned}$$

Así pues, el error de truncación será menor que 0.01 si tomamos como valor aproximado de a a s_{29}.

Nota. En general, las series condicionalmente convergentes convergen lentamente y, por ello, no son muy adecuadas para propósitos de cálculo.

Problemas

1. Pruébese que las siguientes series convergen y estímese a partir de

qué n podemos asegurar que la suma s_n se aproxima a la suma de la serie con un error de truncación menor que el número que se da.

$a)$ $\sum_{k=1}^{\infty} \frac{k}{2^k}$; 10^{-2} $b)$ $\sum_{k=1}^{\infty} \frac{1}{k!}$; 10^{-3}

$c)$ $\sum_{k=1}^{\infty} \frac{2^k}{k!}$; 10^{-3} $d)$ $\sum_{k=1}^{\infty} \frac{1}{k^k}$; 10^{-3}

$e)$ $\sum_{k=1}^{\infty} \frac{1}{k^4}$; 10^{-4} $f)$ $\sum_{k=1}^{\infty} \frac{k+2}{k^5+3k^3}$; 10^{-4}

$g)$ $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{k+3}{k^3+2k}$; 10^{-3} $h)$ $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{k}{2^k}$; 10^{-3}.

5. REORDENACIÓN DE SERIES

Dos propiedades fundamentales de la adición de los números reales son las leyes asociativa y conmutativa: $a+(b+c) = (a+b)+c$ y $a+b = b+a$. Estas propiedades pueden extenderse a cualesquier sumas finitas por *inducción matemática*. Investigaremos ahora su extensión a las series infinitas.

Supongamos que Σa_k es una serie convergente de suma a. Si agrupamos los términos de esta serie dentro de paréntesis para formar nuevos términos, entonces la serie resultante Σb_k también converge a a. Esto es fácil ver ya que la sucesión $\{t_n\}$, donde $t_n = \sum_{k=1}^{n} b_k$, es una subsucesión de $\{s_n\}$ donde $s_n = \sum_{k=1}^{n} a_k$. Por ejemplo, si

$$b_1+b_2+b_3+b_4+\cdots = (a_1+a_2)+a_3+(a_4+a_5+a_6)+a_7+\cdots,$$

entonces $t_1 = s_2$, $t_2 = s_3$, $t_3 = s_6$, $t_4 = s_7$.

Además, si Σa_k diverge a $\pm\infty$, entonces la serie Σb_k divergirá a $\pm\infty$. Sin embargo, si Σa_k oscila, entonces la inserción de paréntesis puede producir una serie convergente. Por ejemplo, la serie

$$\sum_{k=1}^{\infty} (-1)^{k+1} = 1-1+1-1+\cdots$$

oscila. Pero si agrupamos los términos por parejas de modo que

$$\sum b_k = \sum((-1)^{2k-1}+(-1)^{2k}) = (1-1)+(1-1)+\cdots,$$

entonces obtenemos la serie convergente cuyos términos son cero. Mirándolo desde otro punto de vista, si quitamos paréntesis en una serie convergente entonces la nueva serie puede divergir.

5.1 Ejemplo. Pruébese que el decimal periódico $0.51\overline{375}$ es un número racional cuando la barra sobre 375 significa que este es el grupo que se repite.

SOLUCIÓN.

$$0.51\overline{375} = 0.51375375\cdots$$
$$= \frac{5}{10} + \frac{1}{10^2} + \frac{3}{10^3} + \frac{7}{10^4} + \frac{5}{10^5} + \frac{3}{10^6} + \frac{7}{10^7} + \frac{5}{10^8} +$$
$$+ \frac{3}{10^9} + \frac{7}{10^{10}} + \frac{5}{10^{11}} + \cdots .$$

Como esta serie converge $\left(\text{compárese con la serie geométrica } \Sigma \frac{9}{10^k}\right)$, podemos agrupar términos. Entonces

$$0.51\overline{375} = \left(\frac{5}{10} + \frac{1}{10^2}\right) + \left(\frac{3}{10^3} + \frac{7}{10^4} + \frac{5}{10^5}\right) + \left(\frac{3}{10^6} + \frac{7}{10^7} + \frac{5}{10^8}\right) +$$
$$+ \left(\frac{3}{10^9} + \frac{7}{10^{10}} + \frac{5}{10^{11}}\right) + \cdots$$
$$= \frac{51}{10^2} + \frac{375}{10^5} + \frac{375}{10^8} + \frac{375}{10^{11}} + \cdots$$
$$= \frac{51}{10^2} + \frac{375}{10^5}\left(1 + \frac{1}{10^3} + \frac{1}{10^6} + \cdots\right)$$
$$= \frac{51}{10^2} + \frac{375}{10^5} \cdot \frac{10^3}{999}$$
$$= \frac{51\,324}{99\,900}.$$

Consideramos ahora la extensión de la ley conmutativa a las series. ¿Podremos cambiar el orden de los términos en una serie sin afectar la convergencia de la serie? Probaremos que podemos hacerlo si la serie es absolutamente convergente. Cambiando el orden de los términos de una serie se produce una nueva serie que se llama una reordenación de la serie original. Es decir, Σb_k es una *reordenación* de Σa_k si existe una trans-

formación uno-uno f de los enteros positivos sobre los enteros positivos tal que $b_k = a_{f(k)}$.

Estudiaremos primero las series convergentes de términos no negativos.

5.2 Teorema. *Si Σa_k es una serie de términos no negativos que converge a a, entonces cualquier reordenación de Σa_k converge a a.*

PRUEBA. Sea Σb_k una reordenación de Σa_k y sea $s_n = \sum_{k=1}^{n} a_k$ y $t_n = \sum_{k=1}^{n} b_k$. Si a_m es el término de índice mayor en t_n, entonces $t_n \leqslant s_m$. Así pues, para cualquier entero positivo n hay un entero positivo m tal que

$$t_n \leqslant s_m \leqslant a.$$

Si hacemos $b = \sum_{k=1}^{\infty} b_k$ (Σb_k converge ya que $\{t_n\}$ está superiormente acotado por a), entonces $b \leqslant a$. Por otra parte, Σa_k es una reordenación de Σb_k y, por tanto, $a \leqslant b$. Así pues, $b = a$. Y esto completa la prueba.

Para el propósito de extender este resultado a una serie absolutamente convergente Σa_k, escribiremos a_k como la diferencia de dos términos no negativos. Sean

$$a_k^+ = \frac{|a_k| + a_k}{2} \quad \text{y} \quad a_k^- = \frac{|a_k| - a_k}{2}.$$

Entonces, $a_k = a_k^+ - a_k^-$. Como $\pm a_k \leqslant |a_k|$, tenemos

$$0 \leqslant a_k^+ \leqslant |a_k| \quad \text{y} \quad 0 \leqslant a_k^- \leqslant |a_k|.$$

En realidad, si a_k es positivo, entonces a_k^+ es a_k y a_k^- es cero. Si a_k es negativo, entonces a_k^+ es cero y a_k^- es $-a_k$.

5.3 Teorema. *Si Σa_k es una serie absolutamente convergente cuya suma es a, entonces cualquier reordenación de Σa_k converge a a.*

PRUEBA. Consideremos las series Σa_k^+ y Σa_k^- donde a_k^+ y a_k^- están definidas como acabamos de explicar. Estas series son no negativas y los términos de las mismas son menores que o iguales a los términos correspondientes de la serie convergente $\Sigma |a_k|$. Luego, Σa_k^+ y Σa_k^- son series convergentes de términos no negativos y $\Sigma a_k = \Sigma a_k^+ - \Sigma a_k^-$. Sea Σb_k una reordenación de Σa_k. Entonces Σb_k^+ y Σb_k^- son reordenaciones de las series Σa_k^+ y Σa_k^-, respectivamente; por ello, de acuerdo con el teorema 5.2 $\Sigma b_k^+ = \Sigma a_k^+$ y $\Sigma b_k^- = \Sigma a_k^-$. Por tanto,

$$\sum b_k = \sum b_k^+ - \sum b_k^- = \sum a_k^+ - \sum a_k^- = a_k.$$

Y esto completa la prueba.

Si Σa_k es condicionalmente convergente (Σa_k converge, pero $\Sigma |a_k|$ diverge), entonces podemos reordenar Σa_k de manera que se forme una nueva serie Σb_k que converja a cualquier número real, diverja a $\pm\infty$ u oscile. Es decir, si p y q son dos puntos cualesquiera (posiblemente $\pm\infty$) tales que $p \leqslant q$, entonces existe una reordenación Σb_k de Σa_k tal que $\underline{\lim}\ t_n = p$ y $\overline{\lim}\ t_n = q$ donde $t_n = \sum_{k=1}^{n} b_k$. Indicaremos cómo formar esta reordenación Σb_k.

Primero, probamos que si Σa_k es condicionalmente convergente, entonces Σa_k^+ y Σa_k^- divergen. Supongamos que Σa_k^+ converge. Entonces, como $a_k^- = a_k^+ - a_k$, Σa_k^- converge. Luego, $\Sigma |a_k|$ converge ya que $|a_k| = a_k^+ + a_k^-$. Lo que contradice el hecho de que Σa_k es condicionalmente convergente. Luego, Σa_k^+ diverge. De un modo análogo podemos demostrar que Σa_k^- diverge. Así pues, Σa_k^+ y Σa_k^- divergen; en realidad, como ambas son series de términos no negativos, divergen a ∞. Es decir, la serie compuesta de los términos positivos de Σa_k diverge a ∞ y la serie compuesta de los términos negativos de Σa_k diverge a $-\infty$.

Supongamos ahora que p y q $(p \leqslant q)$ son números reales. Formemos la reordenación Σb_k como sigue. Tomemos de manera ordenada exactamente el número de términos positivos de Σa_k necesarios para que la suma obtenida sea mayor que q. Tómense a continuación suficientes términos negativos para que la suma de éstos, junto con la de los términos positivos ya escogidos, sea menor que p. Tómense luego suficientes términos positivos para hacer la suma mayor que q y a continuación bastantes términos negativos para que la suma sea de nuevo menor que p. Si continuamos escogiendo términos de esta manera obtenemos la reordenación Σb_k tal que $\underline{\lim}\ t_n = p$ y $\overline{\lim}\ t_n = q$ donde $t_n = \sum_{k=1}^{n} b_k$. Esta construcción puede efectuarse ya que la serie de términos positivos de Σa_k diverge a ∞ y la serie de términos negativos diverge a $-\infty$. El hecho de que los límites superior e inferior tengan los valores prescritos depende del hecho de que $\lim a_k = 0$.

Si p es un número real, pero q es ∞, entonces modificamos la construcción como sigue. En los pasos donde escogíamos exactamente el número de términos positivos necesarios para hacer que la suma fuera mayor que q, tomamos ahora exactamente los términos positivos necesarios para hacer la suma mayor que $1, 2, \ldots$ sucesivamente. La elección de los términos negativos es la misma que la de la construcción precedente. Análogas modificaciones pueden hacerse si $p = -\infty$.

5.4 Ejemplo. Descríbase la formación de una reordenación de $\sum_{k=1}^{\infty} (-1)^{k+1}\frac{1}{k}$ que converja a cero.

SOLUCIÓN. La serie $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k}$ es condicionalmente convergente y, por tanto, existe una reordenación Σb_k cuya suma es cero. Sea $b_1 = 1$, entonces $t_1 = 1$. Tomemos a continuación los términos negativos suficientes para que la correspondiente suma finita sea negativa:

$$t_2 = 1 - \tfrac{1}{2} = \tfrac{1}{2}$$
$$t_3 = 1 - \tfrac{1}{2} - \tfrac{1}{4} = \tfrac{1}{4}$$
$$t_4 = 1 - \tfrac{1}{2} - \tfrac{1}{4} - \tfrac{1}{6} = \tfrac{1}{12}$$
$$t_5 = 1 - \tfrac{1}{2} - \tfrac{1}{4} - \tfrac{1}{6} - \tfrac{1}{8} = -\tfrac{1}{24}.$$

Tomemos ahora suficientes números positivos para que la suma sea positiva:

$$t_6 = 1 - \tfrac{1}{2} - \tfrac{1}{4} - \tfrac{1}{6} - \tfrac{1}{8} + \tfrac{1}{3} = \tfrac{7}{24}.$$

Para hacer la suma negativa tomamos los términos $-\frac{1}{10}, -\frac{1}{12}, -\frac{1}{14}, -\frac{1}{16}$:

$$t_{10} = 1 - \tfrac{1}{2} - \tfrac{1}{4} - \tfrac{1}{6} - \tfrac{1}{8} + \tfrac{1}{3} - \tfrac{1}{10} - \tfrac{1}{12} - \tfrac{1}{14} - \tfrac{1}{16} = -\tfrac{43}{1680}.$$

Añadiendo otro término positivo se obtiene una suma positiva:

$$t_{11} = 1 - \tfrac{1}{2} - \tfrac{1}{4} - \tfrac{1}{6} - \tfrac{1}{8} + \tfrac{1}{3} - \tfrac{1}{10} - \tfrac{1}{12} - \tfrac{1}{14} - \tfrac{1}{16} + \tfrac{1}{5} = \tfrac{293}{1680}.$$

Si continuamos la elección de términos de esta manera, obtenemos la reordenación Σb_k cuya suma es cero.

Problemas

1. Encuéntrese la expansión decimal de los siguientes números racionales:

a) $\frac{3}{11}$ *b*) $\frac{5}{7}$ *c*) $\frac{9}{13}$.

2. Demuéstrese que cualquier número racional $r = \dfrac{p}{q}$ donde p y q son enteros tiene una expansión decimal periódica.

3. Demuéstrese que los siguientes decimales periódicos son números racionales representándolos en la forma $\dfrac{p}{q}$ donde p y q son enteros.

a) $0.\overline{9}$ *b*) $0.\overline{3}$
c) $0.1\overline{27}$ *d*) $0.32\overline{594}$.

4. Pruébese que todo decimal periódico es un número racional.

5. Descríbase la formación de una reordenación de $\Sigma(-1)^{k+1} \dfrac{1}{k}$ que

a) converja a 1 — *b*) converja a -1
c) diverja a ∞ — *d*) diverja a $-\infty$
e) oscile, con los puntos límites 0 y 1.

6. SERIES DE FUNCIONES

La definición de una serie de funciones es análoga a la de una serie de puntos o números: si $\{\mathbf{f}_k\}$ es una sucesión de funciones de R^p a R^m, entonces la serie $\sum_{k=1}^{\infty} \mathbf{f}_k$ es la sucesión $\{\mathbf{s}_n\}$ donde $\mathbf{s}_n = \sum_{k=1}^{n} \mathbf{f}_k$. Nótese que ahora $\mathbf{s}_n$ es una función: es la función con dominio $\bigcap_{k=1}^{n} \mathscr{D}_{\mathbf{f}_k}$ y regla de correspondencia $\mathbf{s}_n(\mathbf{x}) = \sum_{k=1}^{n} \mathbf{f}_k(\mathbf{x})$. Si para cada $\mathbf{x}$ en un conjunto $\mathscr{E}$, $\sum_{k=1}^{\infty} \mathbf{f}_k(\mathbf{x})$ (es decir, $\{\mathbf{s}_n(\mathbf{x})\}$) converge a un punto $\mathbf{f}(\mathbf{x})$, entonces decimos que la serie $\Sigma \mathbf{f}_k$ *converge puntualmente* a $\mathbf{f}$ sobre $\mathscr{E}$. Por ejemplo, en la introducción a este capítulo probamos que $\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$ para cualquier número real x; expresando esto en términos de funciones, tenemos $\sum_{k=0}^{\infty} \frac{I^k}{k!} = \exp$.

6.1 Ejemplo. Determínese el conjunto sobre el que la serie $\sum_{k=0}^{\infty} I^k$ converge y proporciónese la suma.

Solución. La serie geométrica $\sum_{k=0}^{\infty} x^k = \sum_{k=0}^{\infty} I^k(x)$ converge a $\frac{1}{1-x}$ para $x \in \langle -1, 1\rangle$ y diverge para x fuera de este intervalo. Por tanto, la serie ΣI^k converge puntualmente sobre $\langle -1, 1\rangle$ y la suma de la serie es la función $\frac{1}{1-I}$ con dominio restringido a $\langle -1, 1\rangle$.

6.2 Definición. *La serie* $\sum_{k=1}^{\infty} \mathbf{f}_k$ ***converge uniformemente*** *a* $\mathbf{f}$ *sobre el conjunto* $\mathscr{E}$ *si para cada* $\varepsilon > 0$, *existe un número* N *tal que para todo* $\mathbf{x} \in \mathscr{E}$

$$|\mathbf{s}_n(\mathbf{x}) - \mathbf{f}(\mathbf{x})| < \varepsilon \quad \textit{siempre que } n > N.$$

Es decir, $\Sigma \mathbf{f}_k$ converge uniformemente a $\mathbf{f}$ sobre $\mathscr{E}$ si la sucesión $\{\mathbf{s}_n\}$ converge uniformemente a $\mathbf{f}$ sobre $\mathscr{E}$.

Nota. A causa de la íntima conexión entre la serie de funciones $\Sigma \mathbf{f}_k$ y

las series de valores funcionales $\Sigma \mathbf{f}_k(\mathbf{x})$, frecuentemente hablaremos de las series de valores de la función como si fueran la serie de funciones.

6.3 Ejemplo. Pruébese que $\sum_{k=0}^{\infty} e^{-kx}$ converge uniformemente a $\frac{1}{1-e^{-x}}$ sobre $[1, 5]$.

SOLUCIÓN. Sea un $\varepsilon > 0$ cualquiera y $x \in [1, 5]$.

$$|s_n(x) - f(x)| = \left| \frac{1-e^{-nx}}{1-e^{-x}} - \frac{1}{1-e^{-x}} \right| = \left| \frac{-e^{-nx}}{1-e^{-x}} \right|$$

$$= \frac{e^{-nx}}{1-e^{-x}} \leqslant \frac{e^{-n}}{1-e^{-1}} < 2e^{-n}.$$

Ahora bien, $2e^{-n} < \varepsilon$ si $n > -\ln \frac{\varepsilon}{2}$. Luego si $N = -\ln \frac{\varepsilon}{2}$, para cualquier $x \in [1, 5]$

$$\left| s_n(x) - \frac{1}{1-e^{-x}} \right| < \varepsilon \quad \text{siempre que } n > N.$$

Esto prueba que $\sum_{k=0}^{\infty} e^{-kx}$ converge uniformemente a $\frac{1}{1-e^{-x}}$ sobre $[1, 5]$.

De acuerdo con nuestra discusión sobre las sucesiones de funciones del capítulo 8, sabemos que una serie uniformemente convergente sobre un conjunto $\mathscr{E}$ es puntualmente convergente sobre $\mathscr{E}$, pero que una serie que es puntualmente convergente sobre $\mathscr{E}$ no es necesariamente uniformemente convergente sobre $\mathscr{E}$.

Nótese que si a partir de una sucesión dada $\{\mathbf{s}_n\}$ construimos la serie $\Sigma \mathbf{f}_k$ donde $\mathbf{f}_1 = \mathbf{s}_1$ y $\mathbf{f}_k = \mathbf{s}_k - \mathbf{s}_{k-1}$ para $k > 1$, entonces esta serie tendrá $\{\mathbf{s}_n\}$ como su sucesión de sumas parciales. Así pues, de la sucesión $\{I^n\}$ que es convergente puntualmente, pero no uniformemente convergente sobre $\langle -1, 1]$ obtenemos la serie $I + \sum_{k=2}^{\infty} (I^k - I^{k-1})$ que exhibe el mismo comportamiento.

Hay un criterio para la convergencia uniforme de una serie de funciones que es análogo al criterio de comparación para la convergencia de una serie de puntos. Este criterio, llamado *criterio M de Weierstrass*, se da en el siguiente teorema.

6.4 Teorema. *Si* ΣM_k *converge y si* $|\mathbf{f}_k(\mathbf{x})| \leqslant M_k$ *para todo* $\mathbf{x} \in \mathscr{E}$ *y para toda* k *suficientemente grande, entonces* $\Sigma \mathbf{f}_k$ *converge absolutamente y uniformemente sobre* $\mathscr{E}$.

PRUEBA. De acuerdo con el criterio de comparación (pág. 497) $\Sigma \mathbf{f}_k(\mathbf{x})$ es absolutamente convergente para todo $\mathbf{x} \in \mathscr{E}$. Sean $\mathbf{f}(\mathbf{x}) = \sum_{k=1}^{\infty} \mathbf{f}_k(\mathbf{x})$, $\mathbf{s}_n(\mathbf{x}) = \sum_{k=1}^{n} \mathbf{f}_k(\mathbf{x})$, $M = \sum_{k=1}^{\infty} M_k$, y $t_n \sum_{k=1}^{n} M_k$. Si $|\mathbf{f}_k(\mathbf{x})| \leqslant M_k$ para todo $k > N_1$, entonces para todo $\mathbf{x} \in \mathscr{E}$ y para todos los enteros positivos n y m con $m > n > N_1$ tenemos

$$|\mathbf{s}_m(\mathbf{x}) - \mathbf{s}_n(\mathbf{x})| = \left| \sum_{k=n+1}^{m} \mathbf{f}_k(\mathbf{x}) \right| \leqslant \sum_{k=n+1}^{m} |\mathbf{f}_k(\mathbf{x})| \leqslant \sum_{k=n+1}^{m} M_k = t_m - t_n.$$

Por tanto, $\lim_{m\to\infty} |\mathbf{s}_m(\mathbf{x}) - \mathbf{s}_n(\mathbf{x})| \leqslant \lim_{m\to\infty} (t_m - t_n)$; es decir,

$$|\mathbf{f}(\mathbf{x}) - \mathbf{s}_n(\mathbf{x})| \leqslant M - t_n.$$

Tomemos $\varepsilon > 0$. Como $\lim_{n\to\infty} = M$, existe un número $N > N_1$ tal que $M - t_n < \varepsilon$ siempre que $n > N$. Luego para toda $\mathbf{x} \in \mathscr{E}$

$$|\mathbf{f}(\mathbf{x}) - \mathbf{s}_n(\mathbf{x})| < \varepsilon \quad \text{siempre que } n > N.$$

Lo que demuestra que $\Sigma \mathbf{f}_k$ es uniformemente convergente sobre $\mathscr{E}$.

Ahora probaremos que si una serie de funciones continuas converge uniformemente a una función $\mathbf{f}$ sobre algún conjunto $\mathscr{E}$, entonces $\mathbf{f}$ es continua sobre $\mathscr{E}$.

6.5 Teorema. *Si la serie $\Sigma \mathbf{f}_k$ converge uniformemente sobre el conjunto $\mathscr{E}$ y cada uno de los términos $\mathbf{f}_k$ es una función continua sobre $\mathscr{E}$, entonces la suma de la serie es una función continua sobre $\mathscr{E}$.*

PRUEBA. La convergencia uniforme de la serie $\Sigma \mathbf{f}_k$ a $\mathbf{f}$ es equivalente a la convergencia uniforme de la sucesión $\{\mathbf{s}_n\}$ a $\mathbf{f}$. Como cada uno de los términos $\mathbf{f}_k$ de la serie es una función continua sobre $\mathscr{E}$, de ello resulta que cada uno de los términos $\mathbf{s}_n = \sum_{k=1}^{n} \mathbf{f}_k$ de la sucesión es una función continua sobre $\mathscr{E}$. De donde la continuidad de $\mathbf{f}$ sobre $\mathscr{E}$ sigue del teorema 8.5, pág. 483.

6.6 Ejemplo. Pruébese que la suma de la serie $\sum_{k=1}^{\infty} \frac{I^k}{k^2}$ es una función continua.

SOLUCIÓN. Determinaremos primero el conjunto sobre el que la serie converge, es decir, los valores de x para los que $\Sigma \frac{x^k}{k^2}$ converge. Usando

el criterio de la razón tenemos

$$\lim_{k\to\infty}\left|\frac{x^{k+1}}{(k+1)^2}\Big/\frac{x^k}{k^2}\right| = \lim_{k\to\infty}|x|\left(\frac{k}{k+1}\right)^2 = |x|$$

y, por tanto, la serie converge para $|x| < 1$ y diverge para $|x| > 1$. Si $x = \pm 1$, el criterio de la razón no da información alguna. Sin embargo, sabemos que las series $\Sigma\frac{1}{k^2}$ y $\Sigma\frac{(-1)^k}{k^2}$ convergen ambas. Por tanto, la serie $\Sigma\frac{I^k}{k^2}$ converge sobre el intervalo $[-1, 1]$. Como cada uno de los términos de esta serie es continuo sobre $[-1, 1]$, si podemos probar que la convergencia es uniforme sobre $[-1, 1]$ entonces, según el teorema 6.5, sabremos que la suma es continua sobre $[-1, 1]$. Pero, $\left|\frac{x^k}{k^2}\right| \leqslant \frac{1}{k^2}$ para todo $x\in[-1, 1]$ y para todos los enteros positivos k. Además, $\Sigma\frac{I^k}{k^2}$ converge. Luego de acuerdo con el criterio M de Weierstrass $\Sigma\frac{1}{k^2}$ es uniformemente convergente sobre $[-1, 1]$ y, por tanto, su suma es continua sobre $[-1, 1]$.

Problemas

1. Determínese el conjunto sobre el que cada una de las siguientes series converge.

a) $\sum_{k=1}^{\infty}\frac{I^k}{k}$ *b*) $\sum_{k=1}^{\infty}\frac{I^k}{k^3}$

c) $\sum_{k=1}^{\infty} kx^k$ *d*) $\sum_{k=1}^{\infty}\frac{1}{k}\cos^k$

e) $\sum_{k=1}^{\infty} k^2\exp^k$ *f*) $\sum_{k=1}^{\infty}\frac{k!}{x^k}$

g) $\sum_{k=1}^{\infty}\left(\frac{x}{k}\right)^k$ *h*) $\sum_{k=1}^{\infty}\frac{2^k}{k(k+1)}x^k$.

2. Si $\Sigma\mathbf{f}_k$ es uniformemente convergente sobre un conjunto $\mathscr{E}$ y $\mathscr{F}\subset\mathscr{E}$, pruébese que $\Sigma\mathbf{f}_k$ es uniformemente convergente sobre $\mathscr{F}$.

3. Si $\Sigma\mathbf{f}_k$ y $\Sigma\mathbf{g}_k$ son uniformemente convergentes sobre un conjunto $\mathscr{E}$, pruébese que $\Sigma(\mathbf{f}_k+\mathbf{g}_k)$ es uniformemente convergente sobre $\mathscr{E}$.

4. Pruébese que las siguientes series son uniformemente convergentes sobre los intervalos que se indican.

$a)\ \sum_{k=1}^{\infty} \frac{I^k}{k};\ [-\frac{1}{2}, \frac{1}{2}]$ $\qquad b)\ \sum_{k=1}^{\infty} \frac{I^k}{k^3};\ [-1, 1]$

$c)\ \sum_{k=1}^{\infty} kx^k;\ \langle -\frac{1}{2}, \frac{1}{2} \rangle$ $\qquad d)\ \sum_{k=1}^{\infty} \frac{1}{k} \cos^k;\ \left[\frac{\pi}{4}, \frac{3\pi}{4}\right]$

$e)\ \sum_{k=1}^{\infty} k^2 \exp^k;\ \langle -\infty, -\frac{1}{3}]$ $\qquad f)\ \sum_{k=1}^{\infty} \left(\frac{x}{k}\right)^k;\ [-10, 10].$

5. Pruébese que Σx^k es uniformemente convergente sobre cualquier intervalo $[-a, a]$ donde $0 < a < 1$.

6. Si $\Sigma \mathbf{f}_k$ es una serie de funciones de R^p en R^n y $\mathbf{f}_k = (f_k^1, \ldots, f_k^m)$' pruébese que $\Sigma \mathbf{f}_k$ es uniformemente convergente sobre un conjunto $\mathscr{E}$ si y sólo si cada una de las series componentes Σf_k^j es uniformemente convergente sobre $\mathscr{E}$.

7. Si $\{f_k\}$ es una sucesión de funciones reales que están acotadas y son distintas de cero sobre un conjunto $\mathscr{E}$ y existen números K y r con $r < 1$ tales que

$$\left| \frac{f_{k+1}(x)}{f_k(x)} \right| \leqslant r \quad \text{para todo } k > K \text{ y todo } x \in \mathscr{E},$$

pruébese que Σf_k es uniformemente convergente sobre $\mathscr{E}$.

8. Si $f_k(x) = (-1)^{k+1} \dfrac{1}{k + x^2}$ pruébese que Σf_k es uniformemente convergente sobre $\langle -\infty, \infty \rangle$.

7. INTEGRACIÓN Y DIFERENCIACIÓN DE SERIES

En esta sección, en realidad en el resto del capítulo, consideraremos solamente series de funciones reales de una variable real. Sabemos que para sumas finitas la integral de una suma es la suma de las integrales y la derivada de una suma es la suma de las derivadas. Daremos ahora extensiones de estas propiedades a las series.

7.1 Teorema. *Si la serie Σf_k converge uniformemente a f sobre el intervalo cerrado $[a, b]$ y si cada uno de los términos f_k es integrable sobre $[a, b]$, entonces f es integrable sobre $[a, b]$. Además, si $g_n(x) = \int_a^x f_n$ y $g(x) = \int_a^x f$ para $x \in [a, b]$, entonces Σg_n converge uniformemente a g sobre $[a, b]$.*

PRUEBA. Este teorema es una consecuencia inmediata del teorema correspondiente para sucesiones (teorema 8.7, pág. 484) usando el hecho de que si cada f_k es integrable sobre $[a, b]$ entonces $s_n = \sum_{k=1}^{n} f_k$ es integrable sobre $[a, b]$ e $\int_a^x s_n = \sum_{k=1}^{n} \int_a^x f_k$.

El teorema 7.1 implica que si Σf_k converge uniformemente a f sobre $[a, b]$ y cada uno de los términos f_k es integrable sobre $[a, b]$ entonces f es integrable sobre $[a, b]$ e

$$\int_a^b f = \sum_{k=1}^{\infty} \int_a^b f_k .$$

Este resultado a veces se enuncia como sigue: una serie uniformemente convergente de funciones integrables sobre un intervalo cerrado puede integrarse término a término sobre este intervalo.

7.2 Ejemplo. Pruébese que $\ln \dfrac{1}{1-x} = \sum_{k=1}^{\infty} \dfrac{x^k}{k}$ para $x \in \langle -1, 1 \rangle$.

SOLUCIÓN. Mediante nuestra discusión de las series geométricas supimos que

$$\frac{1}{1-t} = \sum_{k=0}^{\infty} t^k \quad \text{sobre } \langle -1, 1 \rangle.$$

Supongamos que $x \in [0, 1\rangle$. Como para cualquier $t \in [0, x]$, $|t^k| \leqslant x^k$ y Σx^k converge, Σt^k es uniformemente convergente sobre $[0, x]$ según el criterio M de Weierstrass. De donde esta serie puede integrarse término a término sobre el intervalo $[0, x]$ por lo que obtenemos

$$\int_0^x \frac{1}{1-t} = \sum_{k=0}^{\infty} \int_0^x t^k$$

y

$$\ln \frac{1}{1-x} = \sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} = \sum_{k=1}^{\infty} \frac{x^k}{k} .$$

Si $x \in \langle -1, 0 \rangle$, entonces podemos probar que Σt^k es uniformemente convergente sobre $[x, 0]$ de un modo análogo; para cualquier $t \in [x, 0]$, $|t^k| \leqslant |x|^k$ y $\Sigma |x|^k$ converge. Por tanto,

$$\int_x^0 \frac{1}{1-t} = \sum_{k=0}^{\infty} \int_x^0 t^k = -\sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1}$$

y

$$\ln \frac{1}{1-x} = \sum_{k=1}^{\infty} \frac{x^k}{k}.$$

En el ejemplo 7.2, podíamos haber enunciado que $\frac{1}{1-x} = \sum_{k=1}^{\infty} \frac{x^k}{k}$ para $x \in [-1, 1\rangle$. Sin embargo, cuando incluimos el punto -1 en el intervalo, la solución exige algo más que una aplicación directa del teorema 7.1. Si $x \in [-1, 0\rangle$, entonces $\sum_{k=1}^{\infty} \frac{x^k}{k}$ es una serie alternante que converge a, digamos, $f(x)$. Entonces, de acuerdo con 4.5, pág. 510,

$$|f(x) - s_n(x)| < \frac{|x|^{n+1}}{n+1} \leqslant \frac{1}{n+1}.$$

Así pues, para $\varepsilon > 0$ existe un número N $\left(\text{a saber, } N = \frac{1}{\varepsilon}\right)$ tal que para todo $x \in [-1, 0\rangle$

$$|f(x) - s_n(x)| < \varepsilon \quad \text{siempre que } n > N.$$

Esto prueba que $\Sigma \frac{I^k}{k}$ es uniformemente convergente sobre $[-1, 0\rangle$ y, por tanto, su suma es continua sobre $[-1, 0\rangle$. Por ello,

$$f(-1) = \lim_{x \to -1^+} f(x) = \lim_{x \to -1^+} \ln \frac{1}{1-x} = \ln \tfrac{1}{2}.$$

Hemos, pues, probado que

$$\ln \frac{1}{1-x} = \sum_{k=1}^{\infty} \frac{x^k}{k} \quad \text{para } x \in [-1, 1\rangle$$

y, en particular,

$$\ln \tfrac{1}{2} = \sum_{k=1}^{\infty} \frac{(-1)^k}{k}$$

o bien

$$\ln 2 = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}.$$

De acuerdo con el teorema sobre diferenciación de una sucesión de funciones (teorema 8.9, pág. 486) obtenemos el teorema correspondiente sobre diferenciación de una serie de funciones.

7.3 Teorema. *Si* Σf_k *converge a* f *sobre* $[a, b]$, *si cada una de las derivadas* f_k' *es continua sobre* $[a, b]$, *y si* $\Sigma f_k'$ *converge uniformemente sobre* $[a, b]$, *entonces* $f' = \Sigma f_k'$ *sobre* $[a, b]$.

PRUEBA. Como la derivada de una suma es la suma de las derivadas, este teorema es una consecuencia inmediata del teorema 8.9, pág. 486.

7.4 Ejemplo. Pruébese que $\dfrac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1}$ para $x \in \langle -1, 1 \rangle$.

SOLUCIÓN. Sabemos que

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \quad \text{para } x \in \langle -1, 1 \rangle.$$

Consideremos ahora la serie de derivadas: $\sum_{k=1}^{\infty} kx^{k-1}$. Usando el criterio de la razón, tenemos

$$\lim_{k \to \infty} \left| \frac{(k+1)x^k}{kx^{k-1}} \right| = \lim_{k \to \infty} \left(\frac{k+1}{k} \right) |x| = |x|$$

y, por tanto, $\sum_{k=1}^{\infty} kx^{k-1}$ converge para $x \in \langle -1, 1 \rangle$. Tomemos un punto cualquiera $x \in \langle -1, 1 \rangle$ y una a tal que $|x| < a < 1$. Como para cualquier $t \in [-a, a]$, $|kt^{k-1}| \leqslant ka^{k-1}$ y $\sum_{k=1}^{\infty} ka^{k-1}$ converge, el criterio M de Weierstrass nos dice que $\sum_{k=1}^{\infty} kt^{k-1}$ es uniformemente convergente sobre $[-a, a]$. Por tanto, según el teorema 7.3

$$\frac{1}{(1-t)^2} = \sum_{k=1}^{\infty} kt^{k-1} \quad \text{sobre } [-a, a].$$

Así pues, para cada $x \in \langle -1, 1 \rangle$

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1}.$$

Problemas

1. Pruébese que

   a) $\sum_{k=0}^{\infty} (-1)^k x^{2k} = \dfrac{1}{1+x^2}$ sobre $\langle -1, 1 \rangle$.

Sugerencia. $I^k(-x^2) = (-1)^k x^{2k}$.

b) La convergencia de esta serie es uniforme sobre cualquier intervalo $[-a, a]$ donde $0 < a < 1$.

c) $\arctan x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}$ sobre $\langle -1, 1\rangle$

d) $\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}$ es uniformemente convergente sobre $[0, 1]$

e) $\frac{\pi}{4} = \arctan 1 = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}$

2. Pruébese que

a) $\frac{1}{(1-x)^3} = \sum_{k=0}^{\infty} \frac{(k+2)(k+1)}{2} x^k$ sobre $\langle -1, 1\rangle$

b) $\frac{1}{(1-x)^4} = \sum_{k=0}^{\infty} \frac{(k+3)(k+2)(k+1)}{3!} x^k$ sobre $\langle -1, 1\rangle$

c) Si m es un entero positivo cualquiera

$$\frac{1}{(1-x)^m} = \sum_{k=0}^{\infty} \frac{(k+m-1)(k+m-2)\cdots(k+1)}{(m-1)!} x^k \text{ sobre } \langle -1, 1\rangle .$$

3. Verifíquese que

a) $\int_0^1 \sum_{k=1}^{\infty} \frac{1}{k^2+x^2}\, dx = \sum_{k=1}^{\infty} \frac{1}{k} \arctan \frac{1}{k}$

b) $\int_0^{\pi} \sum_{k=1}^{\infty} \frac{\operatorname{sen} kx}{k^2}\, dx = \sum_{k=1}^{\infty} \frac{2}{(2k-1)^3}$

c) $D_x \sum_{k=0}^{\infty} \frac{x^k}{k!} = \sum_{k=0}^{\infty} \frac{x^k}{k!}$, para toda x

d) $D_x \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^{2k-1}}{(2k-1)!} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$, para toda x.

4. Si $\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}$ para $s > 1$, pruébese que

$$\zeta'(s) = -\sum_{k=1}^{\infty} \frac{\ln k}{k^s} \text{ para } s > 1.$$

8. SERIE DE TAYLOR

8.1 Definición. *Si la función f tiene derivadas de todos los órdenes en el punto x_0, entonces la serie* $\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(t-x_0)^k$ *se llama* ***serie de Taylor*** *de f alrededor de x_0.*

8.2 Ejemplo. Proporciónese la fórmula de Taylor de la función seno alrededor de 0.

Solución. Si f = sen, entonces $f^{(1)}$ = cos, $f^{(2)}$ = −sen, $f^{(3)}$ = −cos, $f^{(4)}$ = sen, y, en general, $f^{(4k)}$ = sen, $f^{(4k+1)}$ = cos, $f^{(4k+2)}$ = −sen, $f^{(4k+3)}$ = −cos. Por tanto, $f^{(4k)}(0) = 0$, $f^{(4k+1)}(0) = 1$, $f^{(4k+2)}(0) = 0$, $f^{(4k+3)}(0) = -1$ y la serie de Taylor del seno alrededor de 0 es[1]

$$0+t+0-\frac{1}{3!}t^3+0+\frac{1}{5!}t^5+\cdots=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}t^{2k+1}.$$

Para definir la serie de Taylor de una función f alrededor de algún punto x_0, todo lo que necesitamos es la existencia de todas las derivadas de f en x_0. Sin embargo, es posible que la serie resultante no converja en ningún punto distinto del x_0 e incluso si la serie converge en otros puntos puede ser que no converja al valor que la función f toma en esos puntos (problema 5). Ahora estudiemos bajo qué condiciones la serie de Taylor de una función f converge a f.

Por el teorema de Taylor supimos que si la función f tiene derivadas de todo orden sobre un intervalo $\mathscr{I}$ y $x_0 \in \mathscr{I}$, entonces, para cualquier $x \in \mathscr{I}$ y para cualquier entero positivo n

$$f(x)=\sum_{k=0}^{n-1}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+R_n(x)$$

donde

$$R_n(x)=\frac{1}{(n-1)!}\int_{x_0}^{x} f^{(n)}(t)\,(x-t)^{n-1}\,dt.$$

Es, pues, claro que *la serie de Taylor de f alrededor de x_0 converge a f en un punto $x \in \mathscr{I}$ si y sólo si* $\lim_{n\to\infty} R_n(x) = 0$.

En la determinación del límite de $R_n(x)$ es usualmente conveniente

[1] A la serie de Taylor de una función f alrededor de 0, es habitual llamarla serie de Mac Laurin de f. [N. del T.]

emplear la *forma de Lagrange* para el residuo:

$$R_n(x) = \frac{f^{(n)}(c_n)}{n!}(x-x_0)^n$$

para algún c_n entre x y x_0 si $x \neq x_0$.

8.3 Ejemplo. Pruébese que la serie de Taylor del seno alrededor de 0 converge al seno en toda la recta real.

SOLUCIÓN. Como el seno tiene derivadas de todos los órdenes sobre toda la recta real, tenemos, para cualquier $x \in \mathrm{R}$ y para cualquier entero positivo n,

$$\operatorname{sen} x = \sum_{k=0}^{n-1} \frac{(-1)^k}{(2k+1)!} x^{2k+1} + R_{2n+1}(x)$$

donde $R_{2n+1}(0) = 0$ y $R_{2n+1}(x) \pm \dfrac{\cos c_{2n+1}}{(2n+1)!} x^{2n+1}$ para algún c_{2n+1} entre 0 y x si $x \neq 0$. Así pues, para cualquier $x \in \mathrm{R}$,

$$|R_{2n+1}(x)| \leqslant \frac{|x|^{2n+1}}{(2n+1)!}$$

y, por tanto, $\lim\limits_{n\to\infty} R_n(x) = 0$. Lo que prueba que

$$\operatorname{sen} x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} \quad \text{para cualquier numero real } x,$$

luego,

$$\operatorname{sen} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} I^{2k+1}$$

En algunos casos, es otra forma del residuo, la llama *forma de Cauchy*, la que resulta adecuada para la determinación de $\lim\limits_{n\to\infty} R_n(x)$. La forma de Cauchy puede derivarse de la forma integral del residuo mediante el uso del primer teorema del valor medio para integrales: si $x \neq x_0$

$$R_n(x) = \frac{1}{(n-1)!} \int_{x_0}^{x} f^{(n)}(t)\,(x-t)^{n-1}\,dt$$

$$= \frac{1}{(n-1)!} f^{(n)}(c_n)\,(x-c_n)^{n-1} \int_{x_0}^{x} dt$$

$$= \frac{1}{(n-1)!} f^{(n)}(c_n)\,(x-c_n)^{n-1}\,(x-x_0)$$

donde c_n es un número entre x y x_0.

8.4 Ejemplo. Pruébese que

$$(1+x)^a = 1 + \sum_{k=1}^{\infty} \frac{a(a-1)\cdots(a-k+1)}{k!} x^k$$

para cualquier $x \in \langle -1, 1\rangle$ y para cualquier número real a. Se llama a ésta *serie binomial.*

SOLUCIÓN. Sea $f(x) = (1+x)^a$. Tenemos, entonces:

$$f^{(1)}(x) = a(1+x)^{a-1}, \quad f^{(2)}(x) = a(a-1)(1-x)^{a-2},$$

y, en general,

$$f^{(k)}(x) = a(a-1)\cdots(a-k+1)(1+x)^{a-k}$$

Así pues, la serie de Taylor de f alrededor de 0 es

$$1 + \sum_{k=1}^{\infty} \frac{a(a-1)\cdots(a-k+1)}{k!} I^k.$$

Usando la forma de Cauchy para el residuo, tenemos

$$R_n(x) = \frac{a(a-1)\cdots(a-n+1)}{(n-1)!}(1+c_n)^{a-n}(x-c_n)^{n-1}x$$

donde c_n está entre 0 y x. Sea $c_n = \theta x$, entonces $\theta \in \langle 0, 1\rangle$ y

$$R_n(x) = \frac{a(a-1)\cdots(a-n+1)}{(n-1)!}\left(\frac{1-\theta}{1+\theta x}\right)^{n-1}(1+\theta x)^{a-1}x^n.$$

Supongamos ahora que $x \in \langle -1, 1\rangle$. Entonces $\left|\frac{1-\theta}{1+\theta x}\right|^{n-1} < 1$. Además, si $a > 1$,

$$|1+\theta x|^{a-1} \leqslant (1+|x|)^{a-1}$$

y si $a < 1$,

$$|1+\theta x|^{a-1} \leqslant (1-|x|)^{a-1}$$

Por tanto

$$|R_n(x)| \leqslant (1 \pm |x|)^{a-1} \left|\frac{a(a-1)\cdots(a-n+1)}{(n-1)!}\right| |x^n|$$

y $\lim_{n\to\infty} R_n(x) = 0$ ya que

$$\lim_{n\to\infty} \left| \frac{a(a-1)\cdots(a-n+1)}{(n-1)!} \right| |x|^n = 0 .$$

Esto prueba que

$$(1+x)^a = 1 + \sum_{k=1}^{\infty} \frac{a(a-1)\cdots(a-k+1)}{k!} x^k$$

para cualquier $x \in \langle -1, 1 \rangle$. Si a es un entero positivo, entonces la anterior serie se reduce a una suma finita y lo que tenemos es simplemente el teorema del binomio.

Las funciones que son la suma de su serie de Taylor constituyen una clase importante de funciones y tienen un nombre: funciones analíticas.

8.5 Definición. *La función f es **analítica** en un punto x_0 si hay un intervalo abierto $\mathscr{I}$ que contiene a x_0 tal que f tiene derivadas de todos los órdenes en $\mathscr{I}$ y*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k \quad \textit{para todo } x \in \mathscr{I} .$$

En los ejemplos 8.3 y 8.4 mostramos que la función seno y la función f definida por $f(x) = (1+x)^a$ son analíticas en 0.

Problemas

1. Hállense los primeros cuatro términos de la serie de Taylor de las siguientes funciones alrededor de los puntos que se indican.

a) cos; 0

b) tan; 0

c) sec; 0

d) $\dfrac{I}{1-I^2}$; 0

e) $\dfrac{1}{I}$; 1

f) cos; $\dfrac{\pi}{2}$.

2. Proporciónese la serie de Taylor de la función coseno alrededor de 0 y pruébese que converge a coseno sobre toda la recta real.

3. Proporciónese la serie de Taylor de ln alrededor de 1 y pruébese que converge a ln sobre $\langle 0, 2]$.

Sugerencia: úsese la forma de Cauchy del residuo.

4. Pruébese que

a) $(1+I)^{-\frac{1}{2}} = 1 + \sum_{k=1}^{\infty} \frac{(-\frac{1}{2})(-\frac{3}{2})\cdots(-\frac{1}{2}-k+1)}{k!} I^k$ sobre $\langle -1,1\rangle$

b) $(1-x^2)^{-\frac{1}{2}}$

$$= 1 + \sum_{k=1}^{\infty} (-1)^k \frac{(-\frac{1}{2})(-\frac{3}{2})\cdots(-\frac{1}{2}-k+1)}{k!} x^{2k}, \quad x \in \langle -1, 1\rangle.$$

c) La convergencia de la serie de la parte *b* es uniforme sobre cualquier intervalo $[-a, a]$ donde $0 < a < 1$.

d) arcsen x

$$= x + \sum_{k=1}^{\infty} (-1)^k \frac{(-\frac{1}{2})(-\frac{3}{2})\cdots(-\frac{1}{2}-k+1)}{k!\,(2k+1)} x^{2k+1}, \quad x \in \langle -1, 1\rangle.$$

e) Escríbanse los primeros cuatro términos distintos de cero de la serie de Taylor de arcsen alrededor de 0 y compárense con la serie de parte *d*.

5. Sea f la función definida por $f(0) = 0$ y $f(x) = e^{-(1/x^2)}$, si $x \neq 0$. Proporciónese la serie de Taylor de f alrededor de 0. ¿Para qué valores de x converge la serie de Taylor a $f(x)$?

6. Pruébese que las siguientes funciones son analíticas sobre toda la recta real:

a) exp *b*) sen *c*) cos.

7. Si f y g son analíticas en x_0, pruébese que $f+g$ es analítica en x_0.

9. SERIES DE POTENCIAS

9.1 Definición. *Una serie de la forma* $\sum_{k=0}^{\infty} a_k(I-x_0)^k$ *se llama* ***serie de potencias*** *en* $I-x_0$.

Así pues, la serie de Taylor de una función alrededor del punto x_0 es una serie de potencias en $I-x_0$. En realidad, toda serie de potencias que converge en más de un punto es una serie de Taylor. Probaremos esta afirmación posteriormente.

Consideremos ahora la convergencia de la serie $\sum_{k=0}^{\infty} a_k(x-x_0)^k$. Claramente esta serie converge para $x = x_0$. Usamos el criterio de la raíz para determinar si la serie converge en otros puntos. Si $\overline{\lim}\, |a_k|^{1/k} = \infty$ entonces para $x \neq x_0$

$$\overline{\lim}\, (|a_k|\, |x-x_0|^k)^{1/k} = \infty$$

y, por tanto, $\Sigma a_k(I-x_0)^k$ converge solamente en el punto x_0. Si $\overline{\lim}\ |a_k|^{1/k} < \infty$, entonces

$$\overline{\lim}\ (|a_k|\ |x-x_0|^k)^{1/k} = \overline{\lim}\ |a_k|^{1/k}\ |x-x_0|.$$

Así pues, si $\overline{\lim}\ |a_k|^{1/k} = 0$ entonces $\Sigma a_k(I-x_0)^k$ es absolutamente convergente en toda la recta real. Si $0 < \overline{\lim}\ |a_k|^{1/k} < \infty$, entonces $\Sigma a_k(I-x_0)^k$ es absolutamente convergente cuando $|x-x_0| < \dfrac{1}{\overline{\lim}\ |a_k|^{1/k}}$ y es divergente cuando $|x-x_0| > \dfrac{1}{\overline{\lim}\ |a_k|^{1/k}}$. Enunciaremos en forma de teorema lo que acabamos de probar.

9.2 Teorema. *Si* $\overline{\lim}\ |a_k|^{1/k} = \infty$, *entonces* $\sum_{k=0}^{\infty} a_k(I-x_0)^k$ *converge solamente en el punto* x_0. *Si* $\overline{\lim}\ |a_k|^{1/k} = 0$, *entonces* $\Sigma a_k(I-x_0)^k$ *es absolutamente convergente sobre* $\langle -\infty, \infty\rangle$. *Si* $0 < \overline{\lim}\ |a_k|^{1/k} < \infty$ *y* $r = \dfrac{1}{\overline{\lim}\ |a_k|^{1/k}}$, *entonces* $\Sigma a_k(I-x_0)^k$ *es absolutamente convergente sobre* $\langle x_0-r, x_0+r\rangle$ *y es divergente sobre*

$$\langle -\infty, x_0-r\rangle \cup \langle x_0+r, \infty\rangle.$$

Así pues, si $\overline{\lim}\ |a_k|^{1/k} < \infty$, $\Sigma a_k(I-x_0)^k$ es absolutamente convergente sobre un intervalo abierto: o $\langle -\infty, \infty\rangle$ o $\langle x_0-r, x_0+r\rangle$. Este intervalo se llama el intervalo de convergencia de la serie. Si el intervalo de convergencia es el intervalo finito $\langle x_0-r, x_0+r\rangle$, el teorema 9.2 nada nos dice sobre la convergencia de la serie en los puntos extremos x_0-r y x_0+r. La serie puede convergir o divergir en estos puntos; la convergencia en estos puntos debe investigarse para cada caso particular de serie de potencias que se considere.

9.3 Ejemplo. Determínese el conjunto sobre el que cada una de estas series de potencias converge:

a) $\sum_{k=0}^{\infty} k^k(I-2)^k$ *b)* $\sum_{k=0}^{\infty} \dfrac{1}{(k+1)^k}(I+3)^k$ *c)* $\sum_{k=0}^{\infty} \dfrac{1}{2^k} I^k$

d) $\sum_{k=0}^{\infty} \dfrac{1}{(k+1)2^k} I^k$ *e)* $\sum_{k=0}^{\infty} \dfrac{1}{(k+1)^2\, 2^k} I^k$

SOLUCIÓN

a) Como $\overline{\lim}\ (k^k)^{1/k} = \infty$, $\Sigma k^k(I-2)^k$ converge solamente en el punto 2.

b) Como $\overline{\lim}\left[\frac{1}{(k+1)^k}\right]^{1/k} = 0$, $\Sigma \frac{1}{(k+1)^k}(I+3)^k$ converge sobre $\langle -\infty, \infty\rangle$.

c) Como $\overline{\lim}\left(\frac{1}{2^k}\right)^{1/k} = \frac{1}{2}$, el intervalo de convergencia de $\Sigma \frac{1}{2^k} I^k$ es $\langle -2, 2\rangle$. Investiguemos ahora la convergencia en los puntos -2 y 2. Como

$$\sum_{k=0}^{\infty} \frac{1}{2^k} I^k(-2) = \sum_{k=0}^{\infty} (-1)^k \quad \text{y} \quad \sum_{k=0}^{\infty} \frac{1}{2^k} I^k(2) = \sum_{k=0}^{\infty} 1\,,$$

la serie diverge tanto en uno como en otro puntos. Por tanto, $\Sigma \frac{1}{2^k} I^k$ converge sobre $\langle -2, 2\rangle$.

d) Como $\overline{\lim}\left[\frac{1}{(k+1)2^k}\right]^{1/k} = \frac{1}{2}$, el intervalo de convergencia de $\Sigma \frac{1}{(k+1)2^k} I^k$ es $\langle -2, 2\rangle$. Investigando la convergencia en -2 y 2 tenemos

$$\sum_{k=0}^{\infty} \frac{1}{(k+1)2^k} I^k(-2) = \sum_{k=0}^{\infty} (-1)^k \frac{1}{k+1}$$

y

$$\sum_{k=0}^{\infty} \frac{1}{(k+1)2^k} I^k(2) = \sum_{k=0}^{\infty} \frac{1}{k+1}\,.$$

Por tanto, $\Sigma \frac{1}{(k+1)2^k} I^k$ converge en -2 y diverge en 2 y, por tanto, esta serie converge en $[-2, 2\rangle$.

e) Como $\overline{\lim}\left[\frac{1}{(k+1)^2 2^k}\right]^{1/k} = \frac{1}{2}$, el intervalo de convergencia de $\Sigma \frac{1}{(k+1)^2 2^k} I^k$ es $\langle -2, 2\rangle$. Investigando la convergencia en -2 y 2, tenemos

$$\sum_{k=0}^{\infty} \frac{1}{(k+1)^2 2^k} I^k(-2) = \sum_{k=0}^{\infty} (-1)^k \frac{1}{(k+1)^2}$$

y

$$\sum_{k=0}^{\infty} \frac{1}{(k+1)^2 2^k} I^k(2) = \sum_{k=0}^{\infty} \frac{1}{(k+1)^2}$$

series que, ambas, convergen. Por tanto, $\Sigma \frac{1}{(k+1)^2 2^k} I^k$ converge sobre $[-2, 2]$.

9.4 Teorema. *Si $\mathscr{I}$ es el intervalo de convergencia de $\Sigma a_k(I-x_0)^k$, entonces $\Sigma a_k(I-x_0)^k$ es uniformemente convergente sobre cualquier intervalo cerrado contenido en $\mathscr{I}$.*

PRUEBA. Consideremos un intervalo cerrado $[a, b] \subset \mathscr{I}$. Para todo $x \in [a, b]$,

$$|x-x_0| \leqslant |a-x_0| \quad \text{o} \quad |x-x_0| \leqslant |b-x_0|,$$

según cuál sea el mayor. Supongamos que $|b-x_0|$ es el máximo entre $|a-x_0|$ y $|b-x_0|$. Entonces $|a_k(x-x_0)^k| \leqslant |a_k(b-x_0)^k|$ para toda $x \in [a, b]$ y, como $\Sigma |a_k(b-x_0)^k|$ converge, $\Sigma a_k(I-x_0)^k$ es uniformemente convergente sobre $[a, b]$ según el criterio M de Weierstrass.

De acuerdo con el teorema 9.4, lo anterior nos dice que sobre el intervalo de convergencia la suma de una serie de potencias es continua. Supongamos que el intervalo de convergencia de la serie $\Sigma a_k(I-x_0)^k$ es $\mathscr{I}$. Claramente cada uno de los términos de esta serie es continuo sobre $\mathscr{I}$. Además, si $x \in \mathscr{I}$, entonces existe un intervalo cerrado $[a, b]$ tal que $x \in [a, b] \subset \mathscr{I}$. Por tanto, usando el teorema 6.5, pág. 519, tenemos: *la suma de la serie de potencias $\Sigma a_k(I-x_0)^k$ es continua sobre su intervalo de convergencia.*

Usando el teorema 9.4 y el teorema 7.1, pág. 521, obtenemos: *una serie de potencias puede integrarse término a término sobre cualquier intervalo cerrado contenido en su intervalo de convergencia.* Es decir, si $f = \sum_{k=0}^{\infty} a_k(I-x_0)^k$ y $[a, b]$ está contenido en el intervalo de convergencia de esta serie, entonces

$$\int_a^b f = \sum_{k=0}^{\infty} \int_a^b a_k(I-x_0)^k = \sum_{k=0}^{\infty} a_k \frac{(b-x_0)^{k+1} - (a-x_0)^{k+1}}{k+1}.$$

9.5 Ejemplo. Exprésese $\int_0^x e^{-t^2}\,dt$ como el valor en x de una serie de potencias en I.

SOLUCIÓN. Como

$$\exp = \sum_{k=0}^{\infty} \frac{I^k}{k!} \quad \text{sobre } \langle -\infty, \infty \rangle,$$

$$e^{-t^2} = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{k!} \quad \text{para todo } t \in \langle -\infty, \infty \rangle.$$

Por tanto

$$\int_0^x e^{-t^2}\,dt = \sum_{k=0}^{\infty} \int_0^x \frac{(-1)^k}{k!} t^{2k}\,dt = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)k!}.$$

Probaremos ahora que *una serie de potencias* $\sum_{k=0}^{\infty} a_k(I-x_0)^k$ *puede diferenciarse término a término sobre su intervalo de convergencia.* Demostraremos primero que la serie de las derivadas $\sum_{k=1}^{\infty} ka_k(I-x_0)^{k-1}$ tiene el mismo intervalo de convergencia que la serie original. Esto es cierto, ya que lím $k^{1/k} = 1$ implica que $\overline{\text{lím}}\, |ka_k|^{1/k} = \overline{\text{lím}}\, |a_k|^{1/k}$. Si el intervalo común de convergencia es $\mathscr{I}$, entonces, para cualquier $x \in \mathscr{I}$ hay un intervalo cerrado $[a, b]$ tal que $x \in [a, b] \subset \mathscr{I}$. Según el teorema 9.4, $\sum_{k=1}^{\infty} ka_k(I-x_0)^{k-1}$ es uniformemente convergente sobre $[a, b]$ y, por tanto, según el teorema 7.3, pág. 524, si $f = \sum_{k=1}^{\infty} a_k(I-x_0)^k$ entonces

$$f' = \sum_{k=1}^{\infty} ka_k(I-x_0)^{k-1}.$$

Hemos, pues, probado el siguiente teorema.

9.6 Teorema. *Si una serie de potencias tiene el intervalo de convergencia* $\mathscr{I}$, *entonces la suma de la serie es continua sobre* $\mathscr{I}$, *la serie puede integrarse término a término sobre cualquier intervalo cerrado contenido en* $\mathscr{I}$, *y la serie puede diferenciarse término a término en* $\mathscr{I}$.

El teorema que sigue es una consecuencia inmediata del anterior teorema.

9.7 Teorema. *Si la serie de potencias* $\Sigma a_k(I-x_0)^k$ *tiene el intervalo de convergencia* $\mathscr{I}$ *y su suma es* f, *entonces esta serie es la serie de Taylor de* f *alrededor de* x_0.

PRUEBA. Para probar que $\Sigma a_k(I-x_0)^k$ es la serie de Taylor de f alrededor de x_0, debemos demostrar que $a_k = \dfrac{f^{(k)}(x_0)}{k!}$. Claramente $f(x_0) = a_0$.

Como la serie puede diferenciarse término a término sobre $\mathscr{I}$, es decir, como

$$f' = \sum_{k=1}^{\infty} ka_k(I - x_0)^{k-1}$$

sobre $\mathscr{I}$, tenemos $f'(x_0) = a_1$. La serie de derivadas es también una serie de potencias con intervalo de convergencia $\mathscr{I}$ y, por tanto, puede diferenciarse término a término sobre $\mathscr{I}$:

$$f'' = \sum_{k=2}^{\infty} k(k-1)\, a_k(I - x_0)^{k-2}$$

sobre $\mathscr{I}$. Por tanto, $f''(x_0) = 2!a_2$. Continuando de esta manera podemos demostrar que $a_k = \frac{1}{k} f^{(k)}(x_0)$ para cualquier entero positivo k. (Se puede probar por inducción matemática.) Y esto completa la prueba.

Este teorema prueba que la representación de una función por una serie de potencias en $I - x_0$ es única: cualquier representación de una función por una serie de potencias en $I - x_0$ es la serie de Taylor de la función alrededor de x_0. Así pues, si

$$\sum_{k=0}^{\infty} a_k(I - x_0)^k = \sum_{k=0}^{\infty} b_k(I - x_0)^k$$

en alguna vecindad de x_0, entonces $a_k = b_k$ $(k = 0, 1, \ldots)$. Además, para encontrar la serie de Taylor de una función, no es necesario que los coeficientes se calculen mediante la fórmula $a_k = \frac{1}{k!} f^{(k)}(x_0)$. Frecuentemente hay una forma más fácil de obtener una representación en serie de potencias de una función y no importa cómo se obtenga la serie de potencias, esa serie es la serie de Taylor de f.

9.8 Ejemplo. Obténgase la serie de Taylor alrededor de 0 de la función f definida por $f(x) = \dfrac{1}{3x^2 + 2}$.

SOLUCIÓN. Como

$$\frac{1}{1 - I} = \sum_{k=0}^{\infty} I^k$$

y

$$f(x) = \frac{1}{3x^2 + 2} = \frac{1}{2} \frac{1}{1 - \left(-\frac{3}{2}x^2\right)}$$

tenemos

$$f(x) = \frac{1}{2} \sum_{k=0}^{\infty} \left(-\frac{3}{2} x^2 \right)^k = \sum_{k=0}^{\infty} (-1)^k \frac{3^k}{2^{k+1}} x^{2k}$$

y

$$f = \sum_{k=0}^{\infty} (-1)^k \frac{3^k}{2^{k+1}} I^{2k}$$

Los teoremas 9.6 y 9.7 tienen interpretaciones interesantes en términos de funciones analíticas. El teorema 9.7 implica que si la serie de potencias $\Sigma a_k(I - x_0)^k$ tiene un intervalo de convergencia, entonces la suma de la serie es analítica en x_0. Este teorema también implica que si f es analítica en x_0, entonces f tiene una representación única como una serie de potencias en $I - x_0$. Del teorema 9.6 se deduce que si f es analítica en x_0, entonces todas las derivadas de f son analíticas en x_0.

Problemas

1. Pruébese que si $\text{lím} \left| \frac{a_{k+1}}{a_k} \right| = \infty$, entonces $\Sigma a_k(I - x_0)^k$ converge solamente en el punto x_0. Si $\text{lím} \left| \frac{a_{k+1}}{a_k} \right| = 0$, entonces $\Sigma a_k(I - x_0)^k$ es absolutamente convergente en $\langle -\infty, \infty \rangle$. Si $0 < \text{lím} \left| \frac{a_{k+1}}{a_k} \right| < \infty$ y $r = \text{lím} \left| \frac{a_k}{a_{k+1}} \right|$, entonces $\Sigma a_k(I - x_0)^k$ es absolutamente convergente sobre $\langle x_0 - r, x_0 + r \rangle$ y es divergente sobre $\langle -\infty, x_0 - r \rangle \cup \langle x_0 + r, \infty \rangle$.

2. Determínese el conjunto sobre el que cada una de estas series de potencias converge:

a) $\sum_{k=0}^{\infty} k^2 (I + 1)^k$ *b*) $\sum_{k=0}^{\infty} \frac{1}{k+2} I^k$

c) $\sum_{k=0}^{\infty} k! \, I^k$ *d*) $\sum_{k=0}^{\infty} \frac{1}{k!} I^k$

e) $\sum_{k=1}^{\infty} \frac{k+1}{k^3 + 4k} (I - 2)^k$ *f*) $\sum_{k=0}^{\infty} e^k I^k$.

3. Proporciónese la serie de Taylor de las siguientes funciones alrededor del punto que en cada caso se indica:

a) $f(x) = \frac{1}{x+5}$; 0 *b*) $f(x) = \frac{1}{x+5}$; 1

c) $f(x) = \dfrac{1}{x^2+x-2}$; 0 *d*) senh; 0

e) cosh; 0 *f*) $\begin{cases} f(x) = \dfrac{\operatorname{sen} x}{x}; & 0 \\ f(0) = 1 \end{cases}$

g) $f(x) = \operatorname{sen} x^2$; 0 *h*) $f(x) = a^x$; 0 $(a > 0)$.

4. Proporciónese la serie de Taylor alrededor de 0 para cada una de las siguientes funciones:

a) $\displaystyle\int_0^x \frac{\operatorname{sen} t}{t}\,dt$ *b*) $\displaystyle\int_0^x \operatorname{sen} t^2\,dt$

c) $\displaystyle\int_0^x \frac{1-\cos t}{t}\,dt$ *d*) $\displaystyle\int_0^1 \frac{e^{xt}-e^{-xt}}{t}\,dt$.

5. Pruébese que:

a) $\displaystyle\lim_{x\to 0} \frac{\operatorname{sen} x}{x} = 1$ *b*) $\displaystyle\lim_{x\to 0} \frac{1-\cos x}{x} = 0$

c) $\displaystyle\lim_{t\to 0} \frac{e^{xt}-e^{-xt}}{t} = 2x$ *d*) $\displaystyle\lim_{x\to 0} \frac{a^x-1}{x} = \ln a$

e) $\displaystyle\lim_{(x,y)\to(0,0)} f(x,y) = 0$ donde $f(x,y) = \dfrac{1}{y}\operatorname{sen} xy$, $y \neq 0$ y $f(x,0) = x$.

6. Pruébese que:

a) Si $f = \Sigma a_k I^k$ sobre $\langle -r, r\rangle$ y f es una función par, es decir, si se tiene $f(x) = f(-x)$ para toda $x \in \langle -r, r\rangle$, entonces $a_{2n+1} = 0$ $(n = 0, 1, \ldots)$. Así pues, la serie de potencias es en este caso una serie de potencias pares de I.

b) Si $f = \Sigma a_k I^k$ sobre $\langle -r, r\rangle$ y f es una función impar, es decir, $f(-x) = -f(x)$, entonces $a_{2n} = 0$ $(n = 1, 2, \ldots)$. La serie de potencias es, pues, una serie de potencias impares de I.

7. Encuéntrese la suma de

a) $\displaystyle\sum_{k=1}^{\infty} \frac{k}{3^k}$ *b*) $\displaystyle\sum_{k=1}^{\infty} \frac{k^2}{3^k}$.

10. MULTIPLICACIÓN DE SERIES DE POTENCIAS

La expansión en serie de potencias de un producto fg de funciones puede obtenerse multiplicando la serie de potencias para f por la serie de potencias para g. Esta multiplicación de series de potencias es análoga a la multiplicación de polinomios. Es decir, si $f = \Sigma a_k I^k$ y $g = \Sigma b_k I^k$, entonces

$$\begin{aligned} fg &= (a_0+a_1 I+a_2 I^2+\cdots)(b_0+b_1 I+b_2 I^2+\cdots) \\ &= a_0 b_0+(a_0 b_1+a_1 b_0)I+(a_0 b_2+a_1 b_1+a_2 b_0)I^2+\cdots \\ &= \sum_{k=0}^{\infty} c_k I^k \text{ donde } c_k = \sum_{j=0}^{k} a_j b_{k-j}. \end{aligned}$$

Con el fin de justificar esta multiplicación de series de potencias probaremos el siguiente teorema.

10.1 Teorema. *Si* $\sum_{k=0}^{\infty} a_k$ *y* $\sum_{k=0}^{\infty} b_k$ *son series absolutamente convergentes con sumas a y b respectivamente, entonces* $\sum_{k=0}^{\infty} c_k = ab$ *donde* $c_k = \sum_{j=0}^{k} a_j b_{k-j}$.

PRUEBA. Ordenemos los productos $a_j b_k$ en serie como sigue

10.2 $\quad a_0 b_0+a_0 b_1+a_1 b_1+a_1 b_0+a_0 b_2+a_1 b_2+a_2 b_2+a_2 b_1+a_2 b_0+\ldots$

Consideremos ahora las series cuyos términos son los valores absolutos de los términos de 10.2. Cualquiera de las sumas finitas de esta serie será menor que o igual a

$$\left(\sum_{k=0}^{\infty} |a_k|\right)\left(\sum_{k=0}^{\infty} |b_k|\right)$$

Por tanto, la serie 10.2 es absolutamente convergente. Probemos ahora que la suma de 10.2 es ab. Sea t_n la suma de los primeros n términos de 10.2. Entonces

$$t_{n^2} = \left(\sum_{k=0}^{n-1} a_k\right)\left(\sum_{k=0}^{n-1} b_k\right)$$

y, por tanto, $\lim t_{n^2} = ab$. Como $\{t_n\}$ converge y la subsucesión $\{t_{n^2}\}$ converge a ab, $\lim t_n = ab$. Obsérvese que la serie Σc_k donde $c_k = \sum_{j=0}^{k} a_j b_{k-j}$ puede obtenerse de la 10.2 por reordenación de términos e inserción de paréntesis. Como estas operaciones no afectan la suma de una serie absolutamente convergente, $\Sigma c_k = ab$.

Aplicamos ahora este teorema a la multiplicación de series de potencias.

10.3 Corolario. *Si* $f = \Sigma a_k(I - x_0)^k$ *sobre el intervalo abierto* $\mathscr{I}$ *y* $g = \Sigma b_k(I - x_0)^k$ *sobre el intervalo abierto* $\mathscr{I}'$, *entonces sobre* $\mathscr{I} \cap \mathscr{I}'$

$$fg = \Sigma c_k(I - x_0)^k \text{ donde } c_k = \sum_{j=0}^{k} a_j b_{k-j}.$$

PRUEBA. Si $x \in \mathscr{I} \cap \mathscr{I}'$, entonces $\Sigma a_k(x - x_0)^k$ y $\Sigma b_k(x - x_0)^k$ son absolutamente convergentes. Por tanto, según el teorema 10.1,

$$f(x)\, g(x) = \Sigma c_k(x - x_0)^k \text{ donde } c_k = \sum_{j=0}^{k} a_j b_{k-j}.$$

Y esto completa la prueba.

El corolario 10.3 implica que si f y g son analíticas en x_0, entonces su producto es fg es analítica en x_0.

10.4 Ejemplo. Proporciónense los primeros cuatro términos distintos de cero de la expansión en serie de potencias alrededor de 0 de e^x sen x.

SOLUCIÓN.

$$e^x \operatorname{sen} x = \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots\right)\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right)$$

$$= x + x^2 + \left(-\frac{1}{6} + \frac{1}{2}\right)x^3 + \left(\frac{1}{6} + \frac{1}{6}\right)x^4$$

$$+ \left(\frac{1}{120} - \frac{1}{12} + \frac{1}{24}\right)x^5 + \cdots$$

$$= x + x^2 + \frac{1}{3}x^3 - \frac{1}{30}x^5 + \cdots$$

Esta expansión en serie de potencias es válida sobre toda la recta real.

10.5 Ejemplo. Proporciónese la expansión en serie de potencias alrededor de 0 de sen^2.

SOLUCIÓN. Como $\operatorname{sen} = \sum_{k=0}^{\infty} (-1)^k \dfrac{I^{2k+1}}{(2k+1)!}$ y esta expansión es válida sobre toda la recta real, podemos obtener la serie de potencias para sen^2 por multiplicación y esta expansión será válida para toda la recta real.

$$\operatorname{sen}^2 = \left(\sum_{k=0}^{\infty} (-1)^k \frac{I^{2k+1}}{(2k+1)!}\right)\left(\sum_{k=0}^{\infty} (-1)^k \frac{I^{2k+1}}{(2k+1)!}\right)$$

$$= \sum_{k=0}^{\infty}\left(\sum_{j=0}^{k}\frac{(-1)^j}{(2j+1)!}\,\frac{(-1)^{k-j}}{(2(k-j)+1)!}\right)I^{2k+2}$$

$$= \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+2)!}\left[\sum_{j=0}^{k}\binom{2k+2}{2j+1}\right]I^{2k+2}$$

$$= \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+2)!}\left[\sum_{i=0}^{2k+1}\binom{2k+1}{i}\right]I^{2k+2}$$

$$= \sum_{k=0}^{\infty}(-1)^k\frac{2^{2k+1}}{(2k+2)!}\,I^{2k+2}$$

$$= \sum_{k=1}^{\infty}(-1)^{k-1}\frac{2^{2k-1}}{(2k)!}\,I^{2k}.$$

Esta expansión en serie de potencias puede obtenerse también usando la identidad:

$$\operatorname{sen}^2 x = \frac{1}{2}-\frac{1}{2}\cos 2x = \frac{1}{2}-\frac{1}{2}\sum_{k=0}^{\infty}(-1)^k\frac{(2x)^{2k}}{(2k)!}.$$

Mediante una operación formal de las series de potencias podemos también obtener la expansión en serie de potencias de un cociente de funciones, de la composición de funciones, y de la inversa de una función. Solamente enunciaremos el teorema para cocientes. Los teoremas que justifican estas operaciones de las series de potencias, junto con sus pruebas, pueden encontrarse en los trabajos sobre series infinitas tales como la *Theory and Application of Infinite Series* de K. Knopp (Blackie, Londres, 1951). Sin embargo, tales cuestiones es más conveniente considerarlas en el marco de la teoría de las funciones complejas.

10.6 Teorema. *Si f y g son analíticas en x_0 y $g(x_0) \neq 0$, entonces el cociente f/g es una función analítica en x_0.*

Sea $f(x) = \Sigma a_k(x-x_0)^k$ y $g(x) = \Sigma b_k(x-x_0)^k$. Sabiendo que $f(x)/g(x)$ tiene una expansión en serie de potencias $\Sigma c_k(x-x_0)^k$, podemos determinar los coeficientes c_k mediante multiplicación:

$$\sum_{k=0}^{\infty} a_k(x-x_0)^k = \left[\sum_{k=0}^{\infty} c_k(x-x_0)^k\right]\left[\sum_{k=0}^{\infty} b_k(x-x_0)^k\right]$$

$$= \sum_{k=0}^{\infty}\left(\sum_{j=0}^{k} c_j b_{k-j}\right)(x-x_0)^k.$$

Así pues, $a_k = \sum_{j=0}^{k} c_j b_{k-j}$.

10.7 Ejemplo. Proporciónense los primeros cuatro términos distintos de cero de la expansión en serie de potencias de tan alrededor de 0.

SOLUCIÓN. Como $\tan = \dfrac{\text{sen}}{\cos}$ y $\text{sen} = \sum_{k=0}^{\infty} (-1)^k \dfrac{I^{2k+1}}{(2k+1)!}$ y $\sum_{k=0}^{\infty} (-1)^k \dfrac{I^{2k}}{(2k)!}$, tan tiene una expansión en serie $\sum_{k=0}^{\infty} c_k I^k$ que es válida en alguna vecindad de cero. En realidad, ésta ha de ser una serie de potencias de potencias impares de I ya que tan es una función impar (problema 6, pág. 536). Tenemos, por tanto,

$$I - \frac{I^3}{3!} + \frac{I^5}{5!} - \frac{I^7}{7!} + \cdots = \left(1 - \frac{I^2}{2!} + \frac{I^4}{4!} - \frac{I^6}{6!} + \cdots\right)$$

$$\times\, (c_1 I + c_3 I^3 + c_5 I^5 + c_7 I^7 + \cdots)$$

$$= c_1 I + \left(c_3 - \frac{c_1}{2!}\right) I^3 + \left(c_5 - \frac{c_3}{2!} + \frac{c_1}{4!}\right) I^5$$

$$+ \left(c_7 - \frac{c_5}{2!} + \frac{c_3}{4!} - \frac{c_1}{6!}\right) I^7 + \cdots$$

Así pues, $c_1 = 1$

$$c_3 - \frac{c_1}{2!} = -\frac{1}{3!}; \quad c_3 = \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$$

$$c_5 - \frac{c_3}{2!} + \frac{c_1}{4!} = \frac{1}{5!}; \quad c_5 = \frac{1}{120} + \frac{1}{6} - \frac{1}{24} = \frac{2}{15}$$

$$c_7 - \frac{c_5}{2!} + \frac{c_3}{4!} - \frac{c_1}{6!} = -\frac{1}{7!}; \quad c_7 = \frac{1}{15} - \frac{1}{72} + \frac{1}{720} - \frac{1}{5\,040} = \frac{17}{315}.$$

Por tanto,

$$\tan = I + \frac{1}{3} I^3 + \frac{2}{15} I^5 + \frac{17}{315} I^7 + \cdots$$

Un método equivalente para la obtención de la serie de potencias de tan es mediante el proceso de la división larga. Tenemos

$$\tan x = \frac{x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \dfrac{x^7}{7!} + \cdots}{1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \dfrac{x^6}{6!} + \cdots}.$$

Luego, mediante la división larga,

$$
\begin{array}{r|l}
 & x+\dfrac{x^3}{3}+\dfrac{2}{15}x^5+\dfrac{17}{315}x^7+\cdots \\
\hline
1-\dfrac{x^2}{2}+\dfrac{x^4}{24}-\dfrac{x^6}{720}+\cdots & x-\dfrac{x^3}{6}+\dfrac{x^5}{120}-\dfrac{x^7}{5\,040}+\cdots \\
 & x-\dfrac{x^3}{2}+\dfrac{x^5}{24}-\dfrac{x^7}{720}+\cdots \\
\cline{2-2}
 & \dfrac{x^3}{3}-\dfrac{x^5}{30}+\dfrac{x^7}{840}+\cdots \\
 & \dfrac{x^3}{3}-\dfrac{x^5}{6}+\dfrac{x^7}{72}+\cdots \\
\cline{2-2}
 & \dfrac{2}{15}x^5-\dfrac{4}{315}x^7+\cdots \\
 & \dfrac{2}{15}x^5-\dfrac{x^7}{15}+\cdots \\
\cline{2-2}
 & \dfrac{17}{315}x^7+\cdots \\
 & \dfrac{17}{315}x^7+\cdots \\
\cline{2-2}
\end{array}
$$

Nótese que podemos obtener tantos términos como deseemos de la serie del cociente, pero no tenemos una fórmula para el término general.

Problemas

1. Proporciónense los primeros cuatro términos distintos de cero de las expansiones en series de potencias alrededor de 0 de las siguientes funciones:

a) $f(x)=\dfrac{\operatorname{sen} x}{\sqrt{1-x^2}}$ *b*) $f(x)=\dfrac{\cos x}{x+5}$

c) $f(x)=e^x\cos x$.

2. *a*) Obténgase la expansión en serie de potencias alrededor de 0 de $\cos^2$.

b) Usando series de potencias, pruébese que $\operatorname{sen}^2+\cos^2=1$.

3. Usando series de potencias pruébese que

a) $e^x e^y = e^{x+y}$ *b*) $2 \operatorname{sen} x \cos x = \operatorname{sen} 2x$.

4. Proporciónense los primeros cuatro términos distintos de cero de las expansiones en series de potencias alrededor de 0 de las siguientes funciones:

a) sec *b*) tanh .

11. RESUMEN

En este capítulo presentamos algunos tópicos importantes de la teoría de series. Consideramos primero la convergencia de series de puntos en R^m. Como la convergencia de tales series depende de la convergencia de las series componentes, nos restringimos a la discusión de las series de números reales y desarrollamos unos cuantos criterios de convergencia para estas series. Este tipo de serie es importante para el cálculo. Además, un estudio de la convergencia de series de números reales nos da una base para el estudio de la convergencia de series de funciones.

El concepto de convergencia uniforme cobra importancia cuando estudiamos las propiedades analíticas de las series de funciones. Enfatizamos un tipo particular de series de funciones —las series de potencias— ya que la expansión de una función en una serie de potencias es un método importante para la resolución de muchos problemas. Mediante el uso del teorema de Taylor podemos determinar si una función tiene una representación en serie de potencias; es decir, si una función es analítica. Este teorema también nos provee de un método para encontrar la expansión en serie de potencias de una función analítica, aunque, como hemos visto, puede que haya procedimientos más sencillos de obtener la expansión.

Problemas de repaso

1. Determínense si las siguientes series convergen o divergen.

a) $\sum_{k=1}^{\infty} \frac{k^3+3k}{2^k+2}$ *b*) $\sum_{k=1}^{\infty} (-1)^k \frac{k+3}{2k^2-5}$

c) $\sum_{k=1}^{\infty} \frac{k+3}{2k^2-5}$ *d*) $\sum_{k=1}^{\infty} \frac{3k-2}{k^3+4k}$.

2. Obténgase el valor aproximado de la suma de la serie $\sum_{k=1}^{\infty} \frac{1}{k^3}$ con un error de truncación menor que 1×10^{-2}.

3. Determínense los valores de x para los que la serie $\sum_{k=1}^{\infty} \frac{(x-1)^k}{k3^k}$ converge.

4. Pruébese que $\sum_{k=1}^{\infty} \frac{1}{k^x}$ es uniformemente convergente para $x \in [a, \infty\rangle$ donde $a > 0$.

5. Si Σf_k es uniformemente convergente sobre un conjunto $\mathscr{E}$ y la función g es acotada sobre $\mathscr{E}$, pruébese que $\Sigma g f_k$ es uniformemente convergente sobre $\mathscr{E}$.

6. Sea $\mathscr{S}$ el conjunto de todas las funciones reales acotadas definidas sobre un conjunto $\mathscr{E}$. Si f y g están en $\mathscr{S}$, definamos

$$\|f-g\| = \sup_{x \in \mathscr{E}} |f(x)-g(x)| .$$

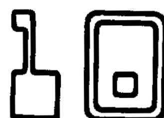
a) Pruébese que

1) $\|f-g\| \geqslant 0$; $\|f-g\| = 0$ implica $f = g$ sobre $\mathscr{E}$
2) $\|f-g\| = \|g-f\|$
3) $\|f-g\| \leqslant \|f-h\| + \|h-g\|$.

b) Si $f_k \in \mathscr{S}$ $(k = 1, 2, \ldots)$ pruébese que Σf_k es uniformemente convergente sobre $\mathscr{E}$ a la función $f \in \mathscr{S}$ si y sólo si para cualquier $\varepsilon > 0$ existe un número N tal que

$$\|f-s_n\| < \varepsilon \quad \text{siempre que } n > N .$$





10 Integrales impropias

1. INTRODUCCIÓN

Recuérdese que la integral definida solamente lo está para funciones acotadas y sobre un intervalo finito. La definición de la integral definida de Riemann no puede aplicarse ni en el caso en que la función no es acotada ni en el caso en el que el intervalo es infinito. Cuando éste es el caso, la definición de la integral se generaliza tomando la integral sobre intervalos finitos adecuados sobre los que la función es acotada y considerando después el límite de estas integrales. Si el límite existe, la integral generalizada se dice que converge y si el límite no existe, se dice que diverge. Tales integrales se llaman impropias o infinitas. Aparte de las integrales impropias, en este capítulo discutiremos las integrales dependientes de un parámetro. Una integral dependiente de un parámetro define una función cuyo dominio

es el conjunto de valores del parámetro para los que la integral está definida. Trataremos cuestiones sobre la continuidad, diferenciabilidad e integrabilidad de tales funciones. En todo este capítulo, las integrales consideradas serán integrales simples (unidimensionales).

2. INTEGRALES IMPROPIAS

Mucho de lo que fue discutido en el capítulo 9 respecto a series infinitas tiene un estrecho paralelo con las integrales impropias (infinitas). Para las integrales impropias tomamos la integral sobre intervalos finitos sobre los que el integrando está acotado y luego consideramos el límite de estas integrales. Para series finitas tomamos sumas finitas —las sumas parciales— y luego consideramos el límite de estas sumas.

Supongamos que f es integrable sobre $[a, b]$ para todo $b \geqslant a$ y sea $F(b) = \int_a^b f$ donde $b \in [a, \infty\rangle$. Entonces $\int_a^\infty f$ se llama *integral impropia (infinita) de primera clase*. Decimos que $\int_a^\infty f$ converge si $\lim\limits_{\infty} F$ existe y en tal caso el valor de $\int_a^\infty f$ es $\lim\limits_{\infty} F$; es decir,

2.1
$$\int_a^\infty f = \lim_{b\to\infty} F(b) = \lim_{b\to\infty} \int_a^b f.$$

Si $\lim\limits_{\infty} F$ no existe, $\int_a^\infty f$ se dice que diverge. El número $F(b) = \int_a^b f$ corresponde a la suma parcial $s_n = \sum\limits_{k=1}^{n} a_k$ de una serie infinita y $\lim\limits_{\infty} F$ corresponde a $\lim\limits_{\infty} s_n$, la suma, de la serie infinita.

2.2 Ejemplo. Evalúese $\int_a^\infty x^n\,dx$, $(a > 0)$.

Solución. Para $n \neq -1$,
$$F(b) = \int_a^b x^n\,dx = \frac{x^{n+1}}{n+1}\Big]_a^b = \frac{1}{n+1}\left[b^{n+1} + a^{n+1}\right].$$

Si $n > -1$, entonces $\lim\limits_{\infty} F = \infty$ e $\int_a^\infty x^n\,dx$ diverge; si $n < -1$, entonces la integral converge y
$$\lim_{\infty} F = \int_a^\infty x^n\,dx = -\frac{a^{n+1}}{n+1}.$$

Para $n = -1$,

$$F(b) = \int_a^b \frac{dx}{x} = \ln b - \ln a$$

y como $\lim_{\infty} F = \infty$, la integral diverge.

Supongamos que f es integrable sobre $[a, b]$ para todo $a \leqslant b$ y sea $F(a) = \int_a^b f$. El valor de la integral impropia $\int_{-\infty}^b f$ es $\lim_{-\infty} F$ si este límite existe. Así pues

2.3
$$\int_{-\infty}^b f = \lim_{a \to -\infty} F(a) = \lim_{a \to -\infty} \int_a^b f.$$

2.4 Ejemplo. Evalúese $\int_{-\infty}^0 e^x \, dx$.

SOLUCIÓN.

$$\int_{-\infty}^0 e^x \, dx = \lim_{a \to -\infty} \int_a^0 e^x \, dx = \lim_{a \to -\infty} e^x \Big]_a^0 = \lim_{a \to -\infty} [1 - e^a] = 1 .$$

Supongamos que f es acotada e integrable sobre todo intervalo $[a, c]$, donde $c \in [a, b\rangle$, pero no acotada sobre $[a, b\rangle$ y sea $F(c) = \int_a^c f$ donde $c \in [a, b\rangle$. Entonces $\int_a^b f$ se llama *integral impropia de segunda clase* y el valor de $\int_a^b f$ es $\lim_{b^-} F$ si es que este límite existe. Así pues

2.5
$$\int_a^b f = \lim_{c \to b^-} F(c) = \lim_{c \to b^-} \int_a^c f.$$

2.6 Ejemplo. Evalúese $\int_a^b \frac{dx}{\sqrt{b-x}}$, donde $a < b$.

SOLUCIÓN.

$$\int_a^b \frac{dx}{\sqrt{b-x}} = \lim_{c \to b^-} \int_a^c \frac{dx}{\sqrt{b-x}} = \lim_{c \to b^-} [-2\sqrt{b-x}]_a^c$$

$$= \lim_{c \to b^-} [-2\sqrt{b-c} + 2\sqrt{b-a}] = 2\sqrt{b-a} .$$

Sea f acotada e integrable sobre todo intervalo $[c, b]$, donde $c \in \langle a, b]$, pero no acotada sobre $\langle a, b]$ y sea $F(c) = \int_c^b f$ donde $c \in \langle a, b]$. Entonces $\int_a^b f$ se dice que converge si $\lim\limits_{a^+} F$ existe, y en tal caso el valor de $\int_a^b f$ es $\lim\limits_{a^+} F$, es decir,

2.7
$$\int_a^b f = \lim_{c \to a^+} F(c) = \lim_{c \to a^+} \int_c^b f.$$

Si $\lim\limits_{a^+} F$ no existe $\int_a^b f$ se dice que diverge.

2.8 Ejemplo. Evalúese $\displaystyle\int_a^b \frac{dx}{(x-a)^{3/2}}$, donde $a < b$.

SOLUCIÓN.

$$\int_a^b \frac{dx}{(x-a)^{3/2}} = \lim_{c \to a^+} \int_c^b \frac{dx}{(x-a)^{3/2}} = \lim_{c \to a^+} \left[-\frac{2}{\sqrt{x-a}} \right]_c^b$$

$$= \lim_{c \to a^+} \left[-\frac{2}{\sqrt{b-a}} + \frac{2}{\sqrt{c-a}} \right] = \infty .$$

Luego la integral diverge.

La integral impropia $\int_{-\infty}^{\infty} f$ se define como $\int_{-\infty}^{a} f + \int_a^{\infty} f$ donde a es un número real cualquiera. Si ambas integrales $\int_{-\infty}^{a} f$ e $\int_a^{\infty} f$ convergen, entonces $\int_{-\infty}^{\infty} f$ se dice que converge y si cualquiera de las dos $\int_{-\infty}^{a} f$ o $\int_a^{\infty} f$ diverge, entonces $\int_{-\infty}^{\infty} f$ se dice que diverge. Podemos probar que $\int_{-\infty}^{\infty} f$ es independiente de la elección de a como sigue : como

$$\int_a^b f = \int_a^c f + \int_c^b f \quad (a < c < b),$$

se sigue que si una de las dos integrales $\int_a^{\infty} f$ o $\int_c^{\infty}$ converge, entonces la otra también converge. Análogamente, $\int_{-\infty}^{a} f$ e $\int_{-\infty}^{c} f$ o ambas convergen

o ambas divergen. Por otra parte, como

$$\int_a^\infty f = \int_a^c f + \int_c^\infty f$$

e

$$\int_{-\infty}^a f = \int_{-\infty}^c f - \int_a^c f,$$

se sigue que

$$\int_{-\infty}^\infty f = \int_{-\infty}^a f + \int_a^\infty f = \int_{-\infty}^c f + \int_c^\infty f$$

y el valor de $\int_{-\infty}^\infty f$ es independiente de la elección del número a.

Finalmente, si la función f tiene infinitas discontinuidades en un número finito de puntos $x_1, \dots, x_n$ donde $a \leqslant x_1 < x_2 < \dots < x_n \leqslant b$, definimos la integral impropia $\int_a^b f$ como sigue

$$\int_a^b f = \int_a^{x_1} f + \int_{x_1}^{y_1} f + \int_{y_1}^{x_2} f + \cdots + \int_{x_{n-1}}^{y_{n-1}} f + \int_{y_{n-1}}^{x_n} f + \int_{x_n}^b f$$

donde y_i es algún punto convenientemente escogido entre x_i y x_{i+1}. Si todas las integrales del segundo miembro de la ecuación anterior convergen, entonces $\int_a^b f$ se dice que converge y tiene un valor igual a la suma de los valores de las integrales de este segundo miembro. Si cualquiera de las integrales de la derecha diverge, entonces se dice que $\int_a^b f$ diverge. Tenemos una extensión obvia si $a = -\infty$ o $b = \infty$.

Toda integral impropia del tipo 2.3, 2.5 o 2.7 es equivalente a una integral impropia de la primera clase.

Las integrales del tipo $\int_{-\infty}^b f$ pueden reducirse a integrales impropias de la primera clase por el cambio de variable $y = -x$:

2.9 $$\int_{-\infty}^b f(x)\,dx = \lim_{a\to\infty} \int_{-a}^b f(x)\,dx = \lim_{a\to\infty} \int_{-b}^a f(-y)\,dy = \int_{-b}^\infty f(-y)\,dy.$$

Si f es continua sobre $[a, b\rangle$, pero no acotada sobre $[a, b\rangle$, entonces la integral impropia de segunda clase $\int_a^b f$, puede reducirse a una integral

impropia de primera clase mediante el cambio de variable $y = \dfrac{1}{b-x}$:

2.10 $$\int_a^b f(x)\,dx = \lim_{c \to b^-} \int_a^c f(x)\,dx = \lim_{c \to b^-} \int_{1/(b-a)}^{1/(b-c)} f\left(b - \frac{1}{y}\right)\frac{dy}{y^2}$$

$$= \int_{1/(b-a)}^{\infty} f\left(b - \frac{1}{y}\right)\frac{dy}{y^2}.$$

Si f es continua sobre $\langle a, b]$, pero no es acotada sobre $\langle a, b]$, entonces la integral impropia $\int_a^b f$ puede reducirse a una integral impropia de primera clase mediante el cambio de variable $y = \dfrac{1}{x-a}$:

211. $$\int_a^b f(x)\,dx = \lim_{c \to a^+} \int_c^b f(x)\,dx = \lim_{c \to a^+} -\int_{1/(c-a)}^{1/(b-a)} f\left(a + \frac{1}{y}\right)\frac{dy}{y^2}$$

$$= \lim_{c \to a^+} \int_{1/(b-a)}^{1/(c-a)} f\left(a + \frac{1}{y}\right)\frac{dy}{y^2} = \int_{1/(b-a)}^{\infty} f\left(a + \frac{1}{y}\right)\frac{dy}{y^2}.$$

Como todo tipo de integral impropia puede, mediante un cambio adecuado de variable, transformarse en una integral impropia de primera clase, enunciaremos y probaremos todos nuestros resultados para este caso. Los resultados para los otros casos pueden darse como corolarios.

Como una simple consecuencia de los teoremas sobre límites en ∞, tenemos el siguiente teorema.

212. Teorema. *Si f y g están acotadas sobre $[a, \infty\rangle$ e $\int_a^\infty f$ e $\int_a^\infty g$ convergen ambas, entonces*

1) $\int_a^\infty (f \pm g)$ *converge e*

$$\int_a^\infty (f \pm g) = \int_a^\infty f \pm \int_a^\infty g;$$

2) $\int_a^\infty cf$ *converge e*

$$\int_a^\infty cf = c\int_a^\infty f$$

para cualquier función constante c.

PRUEBA. Como para toda $b \in [a, \infty\rangle$

$$\int_a^b (f \pm g) = \int_a^b f \pm \int_a^b g,$$

$$\lim_{b\to\infty} \int_a^b (f \pm g) = \lim_{b\to\infty} \int_a^b f \pm \lim_{b\to\infty} \int_a^b g = \int_a^\infty f \pm \int_a^\infty g.$$

Así pues, $\int_a^\infty (f \pm g)$ converge e

$$\int_a^\infty (f \pm g) = \int_a^\infty f \pm \int_a^\infty g.$$

Además, como

$$\lim_{b\to\infty} \int_a^b cf = \lim_{b\to\infty} c \int_a^b f = c \lim_{b\to\infty} \int_a^b f = c \int_a^\infty f,$$

$\int_a^\infty cf$ converge e $\int_a^\infty cf = c \int_a^\infty f$.

La integración por partes es a menudo útil en la evaluación de las integrales impropias. Si f y g tienen derivadas continuas sobre $[a, \infty\rangle$, entonces, para todo $b \in [a, \infty\rangle$

$$\int_a^b fg' = fg\Big]_a^b - \int_a^b f'g.$$

Si se sabe que dos de los tres límites:

$$\lim_{b\to\infty} \int_a^b fg', \quad \lim_{b\to\infty} \int_a^b f'g, \text{ y } \lim_{b\to\infty} [f(b)\,g(b) - f(a)\,g(a)]$$

existen, entonces el tercero también existe y

2.13
$$\int_a^\infty fg' = \lim_{b\to\infty} [fg]_a^b - \int_a^\infty f'g.$$

2.14 Ejemplo. Evalúese $\int_0^\infty xe^{-x}\,dx$.

SOLUCIÓN. Si integramos por partes tomando $f(x) = x$, $f'(x) = 1$, $g'(x) = e^{-x}$, y $g(x) = -e^{-x}$, obtenemos

$$\int_0^\infty xe^{-x}\,dx = \lim_{b\to\infty} \int_0^b xe^{-x}\,dx = \lim_{b\to\infty} \left\{ -xe^{-x}\Big]_0^b + \int_0^b e^{-x}\,dx \right\}$$
$$= \lim_{b\to\infty} -e^{-x}\Big]_0^b = 1.$$

Problemas

1. Evalúense las siguientes integrales impropias:

a) $\displaystyle\int_1^\infty \frac{x^{3/2}+3}{x^3}\,dx$ *b*) $\displaystyle\int_0^\infty xe^{-x^2}\,dx$

c) $\displaystyle\int_2^\infty \frac{dx}{x^4+4x^2}$ *d*) $\displaystyle\int_0^\infty \frac{x^2\,dx}{x^3+1}$

e) $\displaystyle\int_0^2 \frac{2x-1}{(x-2)^3}\,dx$ *f*) $\displaystyle\int_3^\infty \frac{2x-1}{(x-2)^3}\,dx$

g) $\displaystyle\int_0^\infty \frac{x^2\,dx}{(x^2+1)^2}$ *h*) $\displaystyle\int_0^\infty \frac{x\,dx}{1+x^2}$

i) $\displaystyle\int_0^\infty \frac{dx}{x^3+1}$ *j*) $\displaystyle\int_0^\infty e^{-x}\cos x\,dx$

k) $\displaystyle\int_2^\infty \frac{dx}{x\sqrt{x-1}}$ *l*) $\displaystyle\int_1^\infty \frac{\ln x}{x}\,dx.$

2. Evalúense por medio de la integración por partes las siguientes integrales impropias.

a) $\displaystyle\int_0^1 \ln x\,dx$ *b*) $\displaystyle\int_1^2 \frac{x\,dx}{\sqrt{x-1}}$

c) $\displaystyle\int_0^\infty x^2e^{-x}\,dx$ *d*) $\displaystyle\int_0^1 \sqrt{x}\ln x\,dx.$

3. Encuéntrese el área de cada una de las regiones no acotadas que abajo aparecen, si es que tal área existe.

a) $\{(x,y)\,|\,0\leqslant x<\infty,\ 0\leqslant y\leqslant e^{-x}\}$

b) $\left\{(x,y)\,\middle|\,x\in\mathrm{R},\ 0\leqslant y\leqslant \dfrac{1}{x^2+1}\right\}$

c) $\left\{(x,y)\,\middle|\,0\leqslant x<1,\ 0\leqslant y\leqslant \dfrac{1}{\sqrt[3]{1-x}}\right\}$

d) $\left\{(x,y)\,\middle|\,0\leqslant x<4,\ 0\leqslant y\leqslant \dfrac{x^{3/2}}{\sqrt{4-x}}\right\}.$

4. Encuéntrese la longitud de arco de la espiral logarítmica $r = e^{-\theta}$ para $\theta \in [0, \infty\rangle$.

5. Demuéstrese, integrando por partes, que

$$\int_0^\infty x^n e^{-x} dx = n!.$$

6. Pruébese, integrando por partes, que

$$\int_0^\infty x^{2n+1} e^{-x^2} dx = \frac{n!}{2}.$$

7. Pruébese que si F es una función monótona acotada sobre el intervalo $[a, \infty\rangle$, entonces $\lim\limits_{\infty} F$ existe. Si F es no decreciente (no creciente) $\lim\limits_{\infty} F = \sup \{F(x) \mid x \in [a, \infty\rangle\}$ ($\inf \{F(x) \mid x \in [a, \infty\rangle\}$).

3. CRITERIOS DE CONVERGENCIA Y DIVERGENCIA PARA LAS INTEGRALES IMPROPIAS

Como con las series infinitas, necesitamos algunos criterios para la convergencia de la integral $\int_a^\infty f$ que puedan expresarse en términos del integrando f. En esta sección obtendremos algunos criterios de convergencia.

Para una serie infinita Σa_n, tenemos $\lim\limits_{\infty} a_n = 0$ como una condición necesaria de convergencia de la serie (3.1, pág. 497). Para las integrales impropias se tiene un resultado análogo.

3.1 Teorema. *Si f es continua sobre $[a, \infty\rangle$ y $\lim\limits_{\infty} f$ existe, entonces $\lim\limits_{\infty} f = 0$ es una condición necesaria para la convergencia de $\int_a^\infty f$.*

Prueba. Supongamos que $\lim\limits_{\infty} f = L \neq 0$. Si $L > 0$, entonces existe un número $N \geqslant a$ tal que $f(x) > \frac{1}{2}L$ para todo $x > N$. Para todo x_1, b con $b > x_1 > N$, tenemos

$$\int_{x_1}^b f > \int_{x_1}^b \tfrac{1}{2}L = \tfrac{1}{2}L(b-x_1).$$

Como $\lim\limits_{b\to\infty} \frac{1}{2}L(b-a) = \infty$, $\int_{x_1}^\infty f$ también diverge. Ahora bien, $\int_a^b f =$

$\int_a^{x_1} f + \int_{x_1}^b f$ de modo que $\lim_{b\to\infty} \int_{x_1}^b f = \infty$ implica $\lim_{b\to\infty} \int_a^b f = \infty$. La prueba para $L < 0$ es análoga. Las desigualdades están invertidas e $\int_a^\infty f = -\infty$. Por tanto, si $L \neq 0$ entonces $\int_a^\infty f$ diverge.

Sabemos que para las series infinitas $\lim_\infty a_n = 0$ no es condición suficiente para la convergencia de la serie Σa_n. También aquí nos encontramos que $\lim_\infty f = 0$ no es condición suficiente para la convergencia de $\int_a^\infty f$. Para $n \in [-1, 0\rangle$, $\lim_{x\to\infty} x^n = 0$, pero como vimos en el ejemplo 2.2, pág. 546, $\int_a^\infty x^n\,dx$ $(a > 0)$, diverge para $n \geqslant -1$.

La convergencia de la integral $\int_a^\infty f$ no siempre implica que $\lim_\infty f = 0$. Puede ser que el límite no exista. Como se muestra en el siguiente ejemplo, incluso cuando f es continua y no negativa en $[a, \infty\rangle$, $\int_a^\infty f$ puede convergir aunque $\lim_\infty f$ no exista.

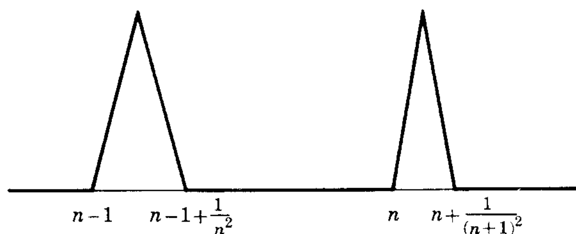


FIGURA 1

3.2 Ejemplo. Pruébese que si

$$f(x) = \begin{cases} 2n^2(x-n+1) & x \in \left\langle n-1,\ n-1+\dfrac{1}{2n^2}\right] \\ -2n^2\left(x-n+1-\dfrac{1}{n^2}\right) & x \in \left\langle n-1+\dfrac{1}{2n^2},\ n-1+\dfrac{1}{n^2}\right] \\ 0 & x \in \left\langle n-1+\dfrac{1}{n^2},\ n\right] \end{cases}$$

para todo entero positivo n, entonces $\int_0^\infty f$ converge.

SOLUCIÓN. (Figura 1.) La función f es continua sobre $[0, \infty\rangle$ y es acotada: $0 \leqslant f(x) \leqslant 1$ para toda $x \in [0, \infty\rangle$. Por otra parte $\lim\limits_{\infty} f$ no existe ya que en toda vecindad de ∞, f toma todos los valores entre 0 y 1 ambos inclusive. Como f es una función no negativa, $F(x) = \int_0^x f$ es no decreciente. El área bajo la gráfica de f sobre el intervalo $[n-1, n]$ es $\frac{1}{2} \cdot \frac{1}{n^2}$. Luego para $x \in [n-1, n]$,

$$F(x) \leqslant \frac{1}{2}\left(1 + \frac{1}{2^2} + \cdots + \frac{1}{n^2}\right) \leqslant \sum_{k=1}^{\infty} \frac{1}{2k^2}$$

y como la serie a la derecha converge, F es acotado. Como F es acotado y no decreciente, de acuerdo con el problema 7, pág. 553, $\lim\limits_{\infty} F = \int_0^{\infty} f$ existe.

El criterio básico para las integrales impropias de funciones no negativas, como para las series infinitas de términos no negativos, es el *criterio de comparación.*

3.3 Teorema. *Si f y g son continuas sobre $[a, \infty\rangle$, $0 \leqslant f(x) \leqslant g(x)$ para toda $x \in [a, \infty\rangle$ e $\int_a^{\infty} g$ converge, entonces $\int_a^{\infty} f$ converge.*

PRUEBA. Sea $F(b) = \int_a^b f$ y $G(b) = \int_a^b g$. Como f y g son funciones no negativas, F y G son no decrecientes y para todo $b \in [a, \infty\rangle$,

$$0 \leqslant F(b) \leqslant G(b) \leqslant \int_a^{\infty} g\,.$$

Así pues F es una función monótona acotada y según el problema 7, pág. 553, $\lim\limits_{\infty} F = \int_a^{\infty} f$ existe.

3.4 Corolario. *Si f y g son continuas sobre $[a, \infty\rangle$, $0 \leqslant g(x) \leqslant f(x)$ para todo $x \in [a, \infty\rangle$, e $\int_a^{\infty} g$ diverge, entonces $\int_a^{\infty} f$ diverge.*

PRUEBA. Supongamos que $\int_a^{\infty} f$ converge. Entonces, según el teorema 3.3, $\int_a^{\infty} g$ converge en contra de la hipótesis. Lo que prueba el corolario.

La forma límite del criterio de comparación es a menudo más fácil de aplicar.

3.5 Corolario. *Supongamos que f y g son continuas sobre $[a, \infty\rangle$ y que $0 \leqslant f(x)$ y $0 < g(x)$ para toda $x \in [a, \infty\rangle$.*

1) *Si* $\lim\limits_{\infty} \dfrac{f}{g} = c \geqslant 0$ *e* $\int_a^\infty g$ *converge, entonces* $\int_a^\infty f$ *converge.*

2) *Si* $\lim\limits_{\infty} \dfrac{f}{g} = c$*, donde* $c > 0$ *o* $c = \infty$*, e* $\int_a^\infty g$ *diverge, entonces* $\int_a^\infty f$ *diverge.*

PRUEBA. Damos la prueba solamente para el caso $c > 0$. Las pruebas para $c = 0$ y $c = \infty$ son análogas (problema 10).

Si $\lim\limits_{\infty} \dfrac{f}{g} = c > 0$, existe un número $N \geqslant a$ tal que $\dfrac{1}{2}c \leqslant \dfrac{f(x)}{g(x)} \leqslant \dfrac{3}{2}c$ o $\dfrac{1}{2}cg(x) \leqslant f(x) \leqslant \dfrac{3}{2}cg(x)$ para todo $x \geqslant N$. Si $\int_a^\infty g$ converge, entonces $\int_N^\infty \dfrac{3}{2}cg$ converge y según el criterio de comparación $\int_N^\infty f$ converge. De donde $\int_a^\infty f$ converge. Si $\int_a^\infty g$ diverge, entonces $\int_N^\infty \dfrac{1}{2}cg$ diverge y según el corolario 3.4, $\int_N^\infty f$ diverge. De donde $\int_a^\infty f$ diverge.

3.6 Ejemplo. ¿Converge la integral $\displaystyle\int_0^\infty \frac{x^3}{2^x}dx$?

SOLUCIÓN. Como para x suficientemente grande 2^x es mucho mayor que x^3, $\dfrac{x^3}{2^x}$ se comporta como $\dfrac{1}{2^x}$. Así pues, podemos intentar aplicar el corolario 3.5 con $f(x) = \dfrac{x^3}{2^x}$ y $g(x) = \dfrac{1}{2^x}$. Sin embargo, en este caso

$$\lim_{\infty} \frac{f}{g} = \lim_{x\to\infty} x^3 = \infty$$

y como $\displaystyle\int_0^\infty \frac{1}{2^x}\,dx$ converge el criterio no puede aplicarse.

Podemos aplicar el corolario 3.5 si tomamos una función un poco mayor

para g. Tomando $g(x) = (\frac{2}{3})^x$, tenemos

$$\lim_{\infty} \frac{f}{g} = \lim_{x\to\infty} \frac{x^3}{2^x(\frac{2}{3})^x} = \lim_{x\to\infty} \frac{x^3}{(\frac{4}{3})^x} = 0.$$

Ahora bien

$$\int_0^\infty g = \int_0^\infty \left(\frac{2}{3}\right)^x dx = \lim_{b\to\infty} \frac{1}{\ln\frac{2}{3}}\left(\frac{2}{3}\right)^x\Big]_0^b = -\frac{1}{\ln\frac{2}{3}}.$$

De donde según el corolario 3.5, $\int_0^\infty \frac{x^3}{2^x}\,dx$ converge.

Como un caso particular del corolario 3.5 obtenemos el *criterio de la potencia.*

3.7 Corolario. *Supongamos que f es continua sobre $[a, \infty\rangle$ donde $a > 0$ y que $0 \leqslant f(x)$ para toda $x \in [a, \infty\rangle$.*

1) *Si $\lim_{x\to\infty} x^r f(x) = c \geqslant 0$ para algún número real $r > 1$, entonces $\int_a^\infty f$ converge.*

2) *Si $\lim_{x\to\infty} x^r f(x) = c$, donde $c > 0$ o $c = \infty$, y $r \leqslant 1$, entonces $\int_a^\infty f$ diverge.*

PRUEBA. Tómese $g(x) = \frac{1}{x^r}$ en el corolario 3.5 y úsese el resultado del ejemplo 2.2, pág. 546.

3.8 Ejemplo. ¿Converge $\int_0^\infty \frac{x^2\,dx}{(a^2+x^2)^{3/2}}$?

SOLUCIÓN. Como $\lim_{x\to\infty} x\frac{x^2}{(a^2+x^2)^{3/2}} = \lim_{x\to\infty} \frac{1}{\left(\frac{a^2}{x^2}+1\right)^{3/2}} = 1 > 0$, la integral dada diverge.

El teorema 3.3 y los corolarios 3.4, 3.5 y 3.7 pueden aplicarse a integrales sobre el intervalo $\langle -\infty, b]$ usando la relación 2.9, pág. 549. Podemos, además, obtener resultados análogos para integrales impropias de segunda clase usando las relaciones (2.10, 2.11) entre las integrales impropias de segunda clase y las de primera clase. Damos ahora el análogo del corolario 3.5 para integrales de segunda clase. Otros resultados aparecen en los problemas.

3.9 *Corolario. Supongamos que f y g son continuas sobre $[a, b\rangle$ y que*

$0 \leqslant f(x)$ *y* $0 < g(x)$ *para todo* $x \in [a, b\rangle$. *Supongamos, además, que* $\lim_{b^-} f = \infty$ *y* $\lim_{b^-} g = \infty$.

1) *Si* $\lim_{b^-} \frac{f}{g} = c \geqslant 0$ *e* $\int_a^b g$ *converge, entonces* $\int_a^b f$ *converge.*

2) *Si* $\lim_{b^-} \frac{f}{g} = c$, *donde* $c > 0$ *o* $c = \infty$, *e* $\int_a^b g$ *diverge, entonces* $\int_a^b f$ *diverge.*

PRUEBA. Mediante el cambio de variable $y = \dfrac{1}{b-x}$, vemos que

$$\lim_{x \to b^-} \frac{f(x)}{g(x)} = \lim_{y \to \infty} \frac{f\left(b - \frac{1}{y}\right)}{g\left(b - \frac{1}{y}\right)} = \lim_{y \to \infty} \frac{f\left(b - \frac{1}{y}\right)\frac{1}{y^2}}{g\left(b - \frac{1}{y}\right)\frac{1}{y^2}}$$

y que $\int_a^b f$ e $\int_a^b g$ son equivalentes a

$$\int_{1/(b-a)}^{\infty} f\left(b - \frac{1}{y}\right)\frac{dy}{y^2} \quad \text{e} \quad \int_{1/(b-a)}^{\infty} g\left(b - \frac{1}{y}\right)\frac{dy}{y^2}$$

respectivamente. El corolario se sigue entonces del corolario 3.5.

3.10 Ejemplo. ¿Converge $\displaystyle\int_0^1 \frac{dx}{\sqrt{(1-x)(2+x)}}$?

SOLUCIÓN. El integrando $f(x) = \dfrac{1}{\sqrt{(1-x)(2+x)}}$ es no negativo y $\lim_{x \to 1^-} f(x) = \infty$. Tomando $g(x) = \dfrac{1}{\sqrt{1-x}}$ tenemos

$$\lim_{x \to 1^-} \frac{f(x)}{g(x)} = \lim_{x \to 1^-} \frac{1}{\sqrt{2+x}} = \frac{1}{\sqrt{3}} > 0$$

y de acuerdo con el corolario 3.9 la integral dada converge si y sólo si $\displaystyle\int_0^1 \frac{dx}{\sqrt{1-x}}$ converge. Ahora bien

$$\int_0^1 \frac{dx}{\sqrt{1-x}} = \lim_{b \to 1^-} \int_0^b \frac{dx}{\sqrt{1-x}} = \lim_{b \to 1^-} \left[-2\sqrt{1-x}\right]_0^b = 2$$

luego la integral dada converge.

Aunque el criterio de comparación es un criterio de convergencia para integrales con integrandos positivos, como con series infinitas, podemos aplicarlo para mostrar también la convergencia de ciertas otras integrales. Dada una integral $\int_a^\infty f$, la integral $\int_a^\infty |f|$ es una integral con integrando no negativo y, por tanto, el criterio de comparación puede aplicarse a $\int_a^\infty |f|$. Probaremos ahora que si $\int_a^\infty |f|$ converge, entonces $\int_a^\infty f$ converge.

3.11 Teorema. *Si* $\int_a^\infty |f|$ *converge, entonces* $\int_a^\infty f$ *converge.*[1]

PRUEBA. Supongamos que $\int_a^\infty |f|$ converge. Como $-|f(x)| \leqslant f(x) \leqslant |f(x)|$, tenemos

$$0 \leqslant f(x) + |f(x)| \leqslant 2|f(x)|.$$

Así pues, según el criterio de comparación (teorema 3.3), $\int_a^\infty (f+|f|)$ converge. De donde

$$\int_a^\infty f = \int_a^\infty (f+|f|-|f|)$$

converge según el teorema 2.12. Y esto completa la prueba.

Decimos que la integral $\int_a^\infty f$ es *absolutamente convergente* si $\int_a^\infty |f|$ converge. Lo que nos dice por tanto el teorema 3.11 es que una integral absolutamente convergente es convergente. Es posible que $\int_a^\infty f$ converja incluso si $\int_a^\infty |f|$ diverge. Si $\int_a^\infty f$ converge, pero $\int_a^\infty |f|$ diverge, entonces decimos que $\int_a^\infty f$ es *condicionalmente convergente.*

De acuerdo con el criterio de la integral para series infinitas si f es una función no creciente y no negativa sobre $[1, \infty\rangle$ entonces $\int_1^\infty f$ y $\Sigma f(k)$ o ambas convergen o ambas divergen. Este criterio puede usarse como un criterio para las integrales cuando la convergencia o divergencia de la serie correspondiente pueda determinarse, por ejemplo, con la ayuda del criterio de la razón o del criterio de la raíz.

[1] El autor supone que $\int_a^b f$ está definida para toda $b \in [a, \infty\rangle$. [N. del T.]

3.12 Ejemplo. Determínese si la integral $\int_1^\infty \frac{\sqrt{x}}{3^x}\,dx$ converge o no.

SOLUCIÓN. El integrando es no negativo y decreciente en $[1, \infty\rangle$. Se satisfacen, pues, las condiciones del criterio de la integral. Ahora bien

$$\lim_{n\to\infty} \frac{f(n+1)}{f(n)} = \lim_{n\to\infty} \frac{\sqrt{n+1}\,3^n}{3^{n+1}\sqrt{n}} = \frac{1}{3}$$

de modo que $\Sigma f(n)$ converge y, por tanto, la integral dada también converge.

Del *criterio para las series alternantes* podemos obtener un criterio que es a veces útil para demostrar la convergencia de una integral en que el integrando va cambiando de signo.

3.13 Teorema. *Sea f continua sobre $[b_1, \infty\rangle$ y sea $\{b_k\}$ una sucesión creciente de ceros de f tal que* lím $b_k = \infty$. *Si f no cambia de signo sobre el intervalo $[b_k, b_{k+1}]$ para toda k y si $\{a_k\}$, donde*

$$a_k = (-1)^{k+1} \int_{b_k}^{b_{k+1}} f,$$

es una sucesión no creciente de términos positivos tales que lím $a_k = 0$, *entonces $\int_{b_1}^\infty f$ converge e $\int_{b_1}^\infty f = \sum_{k=1}^\infty (-1)^{k+1} a_k$.*

PRUEBA. Por el criterio para series alternantes (págs. 505 y 511) sabemos que $\Sigma(-1)^{k+1} a_k$ converge a una suma s y que si $s_n = \sum_{k=1}^n (-1)^{k+1} a_k = \int_{b_1}^{b_{n+1}} f$, entonces $|s - s_n| < a_{n+1}$. Queremos probar que dado un $\varepsilon > 0$ existe un número $M \geqslant b_1$ tal que $\left| s - \int_{b_1}^x f \right| < \varepsilon$ siempre que $x \geqslant M$. Ahora bien, como lím $a_k = 0$, dada $\varepsilon > 0$ existe un entero N tal que $a_{n+1} < \varepsilon/2$ siempre que $n \geqslant N$. Si $x \geqslant M = b_{N+1}$, existe un entero $n \geqslant N$ tal que $x \in [b_{n+1}, b_{n+2}\rangle$. Entonces

$$\left| s - \int_{b_1}^x f \right| = \left| s - \int_{b_1}^{b_{n+1}} f - \int_{b_{n+1}}^x f \right| \leqslant \left| s - \int_{b_1}^{b_{n+1}} f \right| + \left| \int_{b_{n+1}}^x f \right|$$

$$\leqslant |s - s_n| + \left| \int_{b_{n+1}}^{b_{n+2}} f \right| < 2a_{n+1} < \varepsilon.$$

Luego $\lim_{x\to\infty} \int_{b_1}^x f = s$ y la prueba es completa.

3.14 Ejemplo. Pruébese que $\int_0^\infty \frac{\operatorname{sen} x}{x}\, dx$ converge.

SOLUCIÓN. Sea

$$a_k = (-1)^{k+1} \int_{(k-1)\pi}^{k\pi} \frac{\operatorname{sen} x}{x}\, dx .$$

Entonces

$$\begin{aligned} a_{k+1} &= \int_{k\pi}^{(k+1)\pi} \frac{|\operatorname{sen} x|}{x}\, dx = \int_{(k-1)\pi}^{k\pi} \frac{|\operatorname{sen}(y+\pi)|}{y+\pi}\, dy \\ &= \int_{(k-1)\pi}^{k\pi} \frac{|\operatorname{sen} y|}{y+\pi}\, dy \leqslant \int_{(k-1)\pi}^{k\pi} \frac{|\operatorname{sen} x|}{x}\, dx = a_k \end{aligned}$$

y

$$a_k = \int_{(k-1)\pi}^{k\pi} \frac{|\operatorname{sen} x|}{x}\, dx \leqslant \int_{(k-1)\pi}^{k\pi} \frac{dx}{x} \leqslant \frac{1}{k-1} .$$

Luego $\lim a_k = 0$ y la sucesión $\{a_k\}$ satisface las condiciones del teorema 3.13. Por tanto, $\int_0^\infty \frac{\operatorname{sen} x}{x}\, dx$ converge.

Problemas

1. Determínese si las siguientes integrales convergen o divergen.

a) $\displaystyle\int_0^\infty \frac{dx}{\sqrt{1+x}}$ *b)* $\displaystyle\int_0^\infty \frac{dx}{1+x^2}$

c) $\displaystyle\int_1^\infty \frac{dy}{y\sqrt{3y^2+2y+1}}$ *d)* $\displaystyle\int_{-\infty}^\infty \frac{dx}{x^2+4x+5}$

e) $\displaystyle\int_{-\infty}^\infty \frac{dx}{x^2+4x+4}$ *f)* $\displaystyle\int_a^\infty \frac{dx}{\sqrt{x^2-a^2}}$ $(a>0)$

g) $\displaystyle\int_a^\infty \frac{dx}{x\sqrt{x^2-a^2}}$ $(a>0)$ *h)* $\displaystyle\int_0^\infty \frac{\cos x\, dx}{1+\operatorname{sen} x + e^x}$.

2. Establézcase el criterio de la potencia: *sea f continua y no negativa sobre todo intervalo $[a, c]$ donde $a < c < b$, y sea r un número real.*

1) *Si* $\lim\limits_{x \to b^-} (b-x)^r f(x) = c \geqslant 0$ *para* $r < 1$, *entonces* $\int_a^b f$ *converge.*

2) *Si* $\lim\limits_{x \to b^-} (b-x)^r f(x) = c$ *donde* $c > 0$ *o* $c = \infty$, *para* $r \geqslant 1$, *entonces* $\int_a^b f$ *diverge.*

3. Determínese si las siguientes integrales convergen o divergen.

a) $\displaystyle\int_0^1 \frac{\operatorname{sen} x}{x^{3/2}}\, dx$ *b*) $\displaystyle\int_0^2 \frac{dx}{x^2-4}$

c) $\displaystyle\int_0^{\pi/2} \frac{\sqrt{x}}{x + \operatorname{sen} x}\, dx$ *d*) $\displaystyle\int_1^2 \frac{dx}{\sqrt{(x-1)(2-x)}}$.

4. Enúnciense los análogos del teorema 3.3 y del corolario 3.4 para integrales impropias de segunda clase.

5. Pruébese que cada una de las siguientes integrales satisface las condiciones del criterio de la integral y aplíquese el criterio de la razón a las series correspondientes para determinar la convergencia o divergencia.

a) $\displaystyle\int_0^\infty \frac{x^2}{3^x}\, dx$ *b*) $\displaystyle\int_0^\infty x\left(\frac{3}{4}\right)^x dx$

c) $\displaystyle\int_1^\infty \frac{x^3+3x}{5^x+2}\, dx$ *d*) $\displaystyle\int_0^\infty \frac{2^x}{3^x+x}\, dx$.

6. Pruébese el criterio de la raíz *si f es continua sobre* $[a, \infty\rangle$ *y*

$$\lim_{x \to \infty} |f(x)|^{1/x} = L < 1 ,$$

entonces $\int_a^\infty f$ *es absolutamente convergente.*

7. Aplíquese el criterio de la raíz del problema 6 para probar la convergencia de las siguientes integrales.

a) $\displaystyle\int_0^\infty e^{-a^2x^2} \cos bx\, dx \quad (a \pm 0)$ *b*) $\displaystyle\int_0^\infty e^{-a^2x^2} \cosh bx\, dx \quad (a \pm 0)$

c) $\displaystyle\int_0^1 x^{m-1}\left(\ln \frac{1}{x}\right)^n dx \quad (m > 0,\ n > 0)$.

Sugerencia : hágase $\ln \dfrac{1}{x} = u$.

8. Úsese el criterio del teorema 3.13 para establecer la convergencia de las siguientes integrales.

a) $\displaystyle\int_0^\infty xe^{-ax}\,\text{sen}\,x\,dx \quad (a>0)$ *b)* $\displaystyle\int_0^\infty \frac{\cos x}{1+x^2}\,dx$

c) $\displaystyle\int_0^\infty \text{sen}\,x^2\,dx$ *d)* $\displaystyle\int_1^\infty \frac{\text{sen}\,ax\cos bx}{x^2}\,dx.$

9. La gráfica de la ecuación $y^2(a-x)=x^3$ se llama "cisoide de Diocles". Establézcase cuál es la integral para el área entre la cisoide y su asíntota. Demuéstrese que la integral es convergente y evalúese.
Sugerencia: hágase $x = a\cos^2\theta$.

10. Supongamos que f y g son continuas sobre $[a, \infty\rangle$ y que $0 \leqslant f(x)$ y $0 < g(x)$ para todo $x \in [a, \infty\rangle$. Pruébese que:

a) Si $\displaystyle\lim_{\infty} \frac{f}{g} = 0$ e $\displaystyle\int_a^\infty g$ converge, entonces $\displaystyle\int_a^\infty f$ converge.

b) Si $\displaystyle\lim_{\infty} \frac{f}{g} = \infty$ e $\displaystyle\int_a^\infty g$ diverge, entonces $\displaystyle\int_a^\infty f$ diverge.

4. INTEGRALES DEFINIDAS DEPENDIENTES DE UN PARÁMETRO

Sea f una función continua sobre el rectángulo

$$\mathscr{R} = \{(x, y) \mid a \leqslant x \leqslant b,\ c \leqslant y \leqslant d\}$$

y sea F la función definida sobre $[c, d]$ según la regla de correspondencia

$$F(y) = \int_a^b f(x, y)\,dx.$$

Primero demostraremos que F es continua sobre $[c, d]$. La continuidad de F sobre $[c, d]$ implica que F es también integrable sobre $[c, d]$. Si aceptamos la hipótesis adicional de que $D_2 f$ es continua sobre $\mathscr{R}$, probaremos que F es diferenciable sobre $[c, d]$.

4.1 Teorema. *Si f es continua sobre el rectángulo $\mathscr{R} = \{(x, y) \mid a \leqslant x \leqslant b,$ $c \leqslant y \leqslant d\}$, entonces $F(y) = \int_a^b f(x, y)\,dx$ es continua sobre $[c, d]$.*

Prueba. Para todo $y \in [c, d]$ fijo, la función g definida por $g(x) = f(x, y)$

es continua y, por tanto, integrable sobre $[a, b]$. Así pues, F está definida sobre $[c, d]$. Como f es continua sobre el conjunto cerrado y acotado $\mathscr{R}$, f es uniformemente continua sobre $\mathscr{R}$, y para cada $\varepsilon > 0$ existe una $\delta > 0$ tal que

$$|f(x, y+h) - f(x, y)| < \frac{\varepsilon}{b-a}$$

siempre que (x, y) y $(x, y+h)$ pertenecen a $\mathscr{R}$ y $|h| < \delta$. Sea ahora y un punto cualquiera en $[c, d]$. Entonces

$$|F(y+h) - F(y)| = \left| \int_a^b [f(x, y+h) - f(x, y)]\, dx \right|$$

$$\leqslant \int_a^b |f(x, y+h) - f(x, y)|\, dx < \int_a^b \frac{\varepsilon}{b-a}\, dx = \varepsilon$$

siempre que $y+h \in [c, d]$ y $|h| < \delta$. Lo que completa la prueba.

4.2 Teorema. *Si las funciones f y $D_2 f$ son continuas sobre el rectángulo $\mathscr{R} = \{(x, y) \mid a \leqslant x \leqslant b,\ c \leqslant y \leqslant d\}$, entonces $F(y) = \int_a^b f(x, y)\, dx$ es diferenciable sobre $[c, d]$ y*

$$D_y \int_a^b f(x, y)\, dx = F'(y) = \int_a^b D_2 f(x, y)\, dx \quad \textit{sobre } [c, d].$$

PRUEBA. Sea $g(y) = \int_a^b D_2 f(x, y)\, dx$ sobre $[c, d]$. Por el teorema 4.1 supimos que g es continua sobre $[c, d]$ ya que $D_2 f$ es continua sobre $\mathscr{R}$. Por tanto, para $y \in [c, d]$, g es integrable sobre $[c, y]$ e

$$\int_c^y g(u)\, du = \int_c^y \int_a^b D_2 f(x, u)\, dx\, du = \int_a^b \int_c^y D_2 f(x, u)\, du\, dx$$

$$= \int_a^b [f(x, y) - f(x, c)]\, dx = F(y) - F(c).$$

El cambio de orden de integración está justificado ya que $D_2 f$ es continua sobre $\mathscr{R}$ (teorema 9.3 y corolario 9.4, pág. 355). Diferenciando miembros lados de la anterior ecuación, obtenemos por el primer teorema fundamental del cálculo

$$F'(y) = g(y) = \int_a^b D_2 f(x, y)\, dx.$$

4.3 Corolario. (Regla de Leibniz.) *Si f y $D_2 f$ son continuas sobre el rectángulo $\mathscr{R} = \{(x, y) \mid a \leqslant x \leqslant b,\ c \leqslant y \leqslant d\}$ y las funciones g y h son diferenciables sobre $[c, d]$ con $g(y) \in [a, b]$ y $h(y) \in [a, b]$ para todo $y \in [c, d]$, entonces $F(y) = \int_{g(y)}^{h(y)} f(x, y)\,dx$ es diferenciable sobre $[c, d]$ y*

$$D_y \int_{g(y)}^{h(y)} f(x, y)\,dx = F'(y)$$

$$= \int_{g(y)}^{h(y)} D_2 f(x, y)\,dx + f(h(y), y)\,h'(y) - f(g(y), y)\,g'(y).$$

PRUEBA. Sea $G(u, v, y) = \int_u^v f(x, y)\,dx$. Entonces $F(y) = G(g(y), h(y), y)$.

De acuerdo con el primer teorema fundamental del cálculo

$$D_1 G(u, v, y) = \frac{\partial}{\partial u}\left\{-\int_v^u f(x, y)\,dx\right\} = -f(u, y)$$

y

$$D_2 G(u, v, y) = \frac{\partial}{\partial v}\int_u^v f(x, y)\,dx = f(v, y).$$

De acuerdo con el teorema 4.2,

$$D_3 G(u, v, y) = \frac{\partial}{\partial y}\int_u^v f(x, y)\,dx = \int_u^v D_2 f(x, y)\,dx.$$

De donde, según la regla de la cadena,

$$F'(y) = D_1 G(g(y), h(y), y)\,g'(y) + D_2 G(g(y), h(y), y)\,h'(y)$$
$$+ D_3 G(g(y), h(y), y)$$
$$= -f(g(y), y)\,g'(y) + f(h(y), y)\,h'(y) + \int_{g(y)}^{h(y)} D_2 f(x, y)\,dx.$$

Problemas

1. Encuéntrense las derivadas de cada una de las siguientes funciones:

a) $F(y) = \int_0^{\pi} (1 - y \operatorname{sen} x)^2\,dx$ *b)* $F(y) = \int_0^1 \frac{y\,dx}{\sqrt{1 - y^2 x^2}}$

c) $F(y) = \int_{\pi/2}^{\pi} \frac{\operatorname{sen}(xy)}{x}\,dx$ *d)* $F(y) = \int_{\pi/2}^{\pi} \frac{\cos(x^2 y)}{x}\,dx$

e) $F(y) = \int_{-y}^{y} \frac{1-e^{-xy}}{x} dx$ f) $F(y) = \int_{-y}^{y^2} \frac{1-e^{-xy}}{x} dx$

g) $F(y) = \int_{0}^{y^2} \arctan \frac{x}{y^2} dx$ h) $F(y) = \int_{y}^{\cos y} (y^2-x^2)^n dx$;

2. Sea $F(y) = \int_0^y e^{xy} dx$. Evalúese F y luego diferénciese tanto en la forma integral como en la forma evaluada. Obténgase de aquí el valor de $\int_0^y xe^{xy} dx$.

3. Pruébese que si $w(t)$ satisface la ecuación diferencial con coeficientes constantes:

$$L[x(t)] = x^{(n)}(t)+a_1 x^{(n-1)}(t)+ \cdots + a_{n-1} x'(t)+a_n x(t) = 0$$

y las condiciones iniciales:

$$w(0) = w'(0) = \cdots = w^{(n-2)}(0) = 0, \; w^{(n-1)}(0) = 1 ,$$

entonces

$$x(t) = \int_0^t w(t-s) f(s) ds$$

satisface la ecuación diferencial $L[x(t)] = f(t)$.

4. Sea $F(y) = \int_0^1 \frac{x^y-1}{\ln x} dx$. Encuéntrese $F'(y)$ y evalúese la integral. Partiendo de $F'(y)$ evalúese $F(y)$. ¿Cuál es el dominio de F?

5. Sea $F(y) = \int_0^y \frac{(y-x)^{n-1}}{(n-1)!} f(x) dx$. Pruébese que $F(0) = F'(0) = \cdots =$ $F^{n-1}(0) = 0$ y $F^{(n)}(y) = f(y)$. Conclúyase de aquí que

$$\int_0^y \int_0^{y_{n-1}} \cdots \int_0^{y_2} \int_0^{y_1} f(x) dx\, dy_1 \cdots dy_{n-2}\, dy_{n-1} = \int_0^y \frac{(y-x)^{n-1}}{(n-1)!} f(x) dx .$$

6. La función de Bessel J_0 puede definirse por la regla de correspondencia

$$J_0(x) = \frac{1}{\pi} \int_{-1}^{1} \frac{\cos xt}{\sqrt{1-t^2}} dt .$$

Pruébese que J_0 satisface la ecuación diferencial (ecuación de Bessel)

$$J_0''(x) + \frac{1}{x} J_0'(x) + J_0(x) = 0.$$

Sugerencia: intégrese J_0' por partes.

7. Supongamos que y satisface la ecuación integral

$$y(x) = 26x - 10 + 6 \int_0^1 (t-x)\, y(t)\, dt.$$

Encuéntrese la ecuación diferencial satisfecha por y y resuélvase para y.

8. Supongamos que y satisface la ecuación integral

$$y(x) = 1 - 2x - 4x^2 + \int_0^x [3 + 6(x-t) - 4(x-t)^2]\, y(t)\, dt.$$

Encuéntrese la ecuación diferencial satisfecha por y. ¿Cuáles son las condiciones iniciales $y(0)$, $y'(0)$ y $y''(0)$?

9. Resuélvase la ecuación integral

$$y(x) = a \operatorname{sen} x + \int_0^x \operatorname{sen}(x-t)\, y(t)\, dt.$$

10. La integral elíptica incompleta de primera clase se define por la regla de correspondencia

$$F(z, k) = \int_0^z \frac{d\theta}{(1 - k^2 \operatorname{sen}^2 \theta)^{1/2}}, \quad 0 \leqslant k \leqslant 1.$$

Pruébese que:

a) $F(z, k)$ crece con k para $z > 0$

b) $\lim_{k \to 1} F(\frac{1}{2}\pi, k) = \infty$.

5. INTEGRALES IMPROPIAS DEPENDIENTES DE UN PARÁMETRO

En la sección previa las integrales consideradas eran integrales definidas. En esta sección extenderemos los resultados obtenidos allí a las integrales impropias. Recuérdese que cuando consideramos cuestiones de continuidad, integrabilidad y diferenciabilidad de series infinitas de funciones, se introdujo el concepto de convergencia uniforme para obtener condiciones suficientes que nos asegurasen estas propiedades. De nuevo introduciremos este

concepto de convergencia uniforme en las integrales impropias para proveernos de condiciones suficientes y asegurarnos la continuidad, integrabilidad y diferenciabilidad de las integrales impropias dependientes de un parámetro.

5.1 Definición. *Sea E un conjunto de números reales y sea f una función de R^2 en R definida sobre $[a, \infty\rangle \times \mathscr{E}$. La integral* $\int_a^\infty f$ *se dice que es* ***uniformemente convergente a*** F ***sobre el conjunto*** $\mathscr{E}$ *si para cada $\varepsilon > 0$ existe un número N tal que para todo $y \in \mathscr{E}$*

$$\left| F(y) - \int_a^b f(x, y)\,dx \right| < \varepsilon \quad \textit{siempre que } b > N.$$

5.2 Ejemplo. Supongamos que $\mathscr{E} = [1, 5]$ y $f(x, y) = e^{-xy}$. Pruébese que $\int_0^\infty f$ converge uniformemente a I^{-1} sobre $[1, 5]$.

Solución. Tomemos una $\varepsilon > 0$. Si $y \in [1, 5]$, entonces

$$\left| \frac{1}{y} - \int_0^b e^{-xy}\,dx \right| = \left| \frac{1}{y} + \frac{1}{y} e^{-by} - \frac{1}{y} \right| = \frac{1}{y} e^{-by} \leqslant e^{-b}.$$

Ahora bien, $e^{-b} < \varepsilon$ si $-b < \ln \varepsilon$, es decir, si $b > -\ln \varepsilon$. Así pues, si $N = \ln \frac{1}{\varepsilon}$, entonces

$$\left| \frac{1}{y} - \int_0^b e^{-xy}\,dx \right| < \varepsilon \quad \text{siempre que } b > N.$$

Probaremos ahora que si la integral de una función continua de dos variables converge uniformemente a una función F sobre un intervalo $[c, d]$, entonces F es continua sobre $[c, d]$. Este teorema es el análogo del teorema 6.5, pág. 519, sobre la continuidad de una serie uniformemente de funciones continuas. Generaliza el resultado del teorema 4.1 a las funciones impropias.

5.3 Teorema. *Si f es continua sobre la franja $\mathscr{R} = \{(x, y) \mid x \in [a, \infty\rangle$ $y \in [c, d]\}$ y $F(y) = \int_a^\infty f(x, y)\,dx$ converge uniformemente sobre $[c, d]$, entonces F es continua sobre $[c, d]$.*

Prueba. Sea y un punto cualquiera de $[c, d]$ y sea $\varepsilon > 0$. Deseamos probar

que hay un número $\delta > 0$ tal que

$$|F(y+h) - F(y)| < \varepsilon$$

siempre que $y+h \in [c, d]$ y $|h| < \delta$. Como $\int_a^\infty f(x, y)\,dx$ converge uniformemente sobre $[c, d]$, existe un número b tal que para todo $y \in [c, d]$

$$\left| F(y) - \int_a^b f(x, y)\,dx \right| < \frac{\varepsilon}{3}.$$

Según el teorema 4.1, pág. 563, $\int_a^b f(x, y)\,dx$ es continua sobre $[c, d]$. Así pues, existe una $\delta > 0$ tal que

$$\left| \int_a^b [f(x, y+h) - f(x, y)]\,dx \right| < \varepsilon/3$$

siempre que $y+h \in [c, d]$ y $|h| < \delta$. Por tanto, para cualquier $y+h \in [c, d]$ tal que $|h| < \delta$ tenemos

$$\left| F(y+h) - F(y) \right| \leqslant \left| F(y+h) - \int_a^b f(x, y+h)\,dx \right|$$
$$+ \left| \int_a^b [f(x, y+h) - f(x, y)]\,dx \right| + \left| \int_a^b f(x, y)\,dx - F(y) \right|$$
$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon\,.$$

Y esto completa la prueba.

La prueba más sencilla para la convergencia uniforme de las integrales infinitas es la correspondiente al criterio M de Weierstrass para la convergencia uniforme de las series infinitas.

5.4 Teorema. *Si $\int_a^\infty M(x)\,dx$ converge y si $|f(x, y)| \leqslant M(x)$ para toda $y \in [c, d]$ y para toda x suficientemente grande, entonces $\int_a^\infty f(x, y)\,dx$ absoluta y uniformemente sobre $[c, d]$.*

PRUEBA. Por el criterio de comparación (teorema 3.3, pág. 555), $\int_a^\infty f(x, y)\,dx$ es absolutamente convergente para todo $y \in [c, d]$. Sea $F(y) = \int_a^\infty f(x, y)\,dx$. Si $|f(x, y)| \leqslant M(x)$ para todo $x > N_1$ y para

todo $y \in [c, d]$, entonces para todo $y \in [c, d]$ y para todos los números b_1 y b_2 tales que $b_2 > b_1 > N_1$ tenemos

$$\left| \int_a^{b_2} f(x, y)\,dx - \int_a^{b_1} f(x, y)\,dx \right| = \left| \int_{b_1}^{b_2} f(x, y)\,dx \right| \leqslant \int_{b_1}^{b_2} |f(x, y)|\,dx$$
$$\leqslant \int_{b_1}^{b_2} M(x)\,dx\,.$$

Por tanto, $\lim_{b_2 \to \infty} \left| \int_a^{b_2} f(x, y)\,dx - \int_a^{b_1} f(x, y)\,dx \right| \leqslant \lim_{b_2 \to \infty} \int_{b_1}^{b_2} M(x)\,dx$;
es decir,

$$\left| F(y) - \int_a^{b_1} f(x, y)\,dx \right| \leqslant \int_{b_1}^{\infty} M(x)\,dx.$$

Tómese un $\varepsilon > 0$ cualquiera. Como $\int_a^{\infty} M(x)\,dx$ converge, existe un número $N \geqslant N_1$ tal que $\int_{b_1}^{\infty} M(x)\,dx < \varepsilon$ siempre que $b_1 > N$. Luego para todo $y \in [c, d]$

$$\left| F(y) - \int_a^{b_1} f(x, y)\,dx \right| < \varepsilon \quad \text{siempre que } b_1 > N\,.$$

Lo que prueba que $\int_a^{\infty} f(x, y)\,dx$ es uniformemente convergente sobre $[c, d]$.

5.5 Ejemplo. La "función gamma" está definida por la integral impropia

$$\Gamma(y) = \int_0^{\infty} x^{y-1} e^{-x}\,dx \quad (y > 0)\,.$$

Pruébese que la función gamma es continua sobre $\langle 0, \infty \rangle$.

SOLUCIÓN. Probaremos primero que $F_1(y) = \int_1^{\infty} x^{y-1} e^{-x}\,dx$ es uniformemente convergente sobre $[0, d]$ para todo $d > 0$. Luego F_1 es continua sobre $[0, d]$ para todo $d > 0$. Por tanto, F_1 es continua sobre $[0, d]$ y como d es un número positivo arbitrario, F_1 es continua sobre $[0, \infty\rangle$. Ahora bien,

$$\lim_{x \to \infty} x^2 x^{d-1} e^{-x} = \lim_{x \to \infty} x^{d+1} e^{-x} = 0$$

y, por tanto, según el criterio de la potencia (corolario 3.7, pág. 557) $\int_1^{\infty} x^{d-1} e^{-x}\,dx$ converge. Como $0 \leqslant x^{y-1} e^{-x} \leqslant x^{d-1} e^{-x}$ para todo

$x \in [1, \infty\rangle$ y para toda $y \in [0, d]$, $F_1(y)$ converge uniformemente sobre $[0, d]$ según el criterio M de Weierstrass.

Para $y \geqslant 1$, $F_2(y) = \int_0^1 x^{y-1} e^{-x}\, dx$ es una integral definida y la continuidad de F_2 sobre $[1, \infty\rangle$ se sigue del teorema 4.1, pág. 563. Para $y < 1$, $F_2(y)$ es una integral impropia. Demostraremos a continuación que $F_2(y)$ converge uniformemente sobre $[c, 1]$ para todo $c \in \langle 0, 1\rangle$. Luego F_2 es continua sobre $[c, 1]$ y como c es un número arbitrario de $\langle 0, 1\rangle$, F_2 es continua sobre $[0, 1]$. Sea $c \in \langle 0, 1\rangle$. Entonces

$$0 \leqslant x^{y-1} e^{-x} \leqslant x^{c-1}\, e^{-x} \leqslant x^{c-1}$$

para todo $y \in [c, 1]$ y todo $x \in \langle 0, 1]$. Como $\int_0^1 x^{c-1}\, dx$ converge para $c > 0$, $F_2(y)$ converge uniformemente sobre $[c, 1]$ según el criterio M de Weierstrass.

Hemos probado que F_1 es continua sobre $[0, \infty\rangle$ y F_2 es continua sobre $\langle 0, \infty\rangle$. Por tanto $\Gamma = F_1 + F_2$ es continua sobre $\langle 0, \infty\rangle$.

Sabemos que para integrales definidas dependientes de un parámetro si f es continua sobre el rectángulo $\mathscr{R} = \{(x, y) \mid a \leqslant x \leqslant b,\ c \leqslant y \leqslant d\}$ y si $F(y) = \int_a^b f(x, y)\, dx$, entonces $\int_c^d F(y)\, dy = \int_c^d \int_a^b f(x, y)\, dx\, dy = \int_a^b \int_c^d f(x, y)\, dy\, dx$ —el orden de integración puede intercambiarse. Además, si f y $D_2 f$ son continuas sobre el rectángulo $\mathscr{R}$, entonces $F'(y) = \int_a^b D_2 f(x, y)\, dx$ —el orden de integración y diferenciación puede intercambiarse. Extenderemos ahora estos resultados a las integrales impropias. Para las integrales impropias estas operaciones corresponden a la integración y diferenciación término a término de las series.

5.6 Teorema. *Si la función f es continua sobre la franja $\mathscr{R} = \{(x, y) \mid x \in [a, \infty\rangle,\ y \in [c, d]\}$ y $F(y) = \int_a^\infty f(x, y)\, dx$ converge uniformemente sobre $[c, d]$, entonces*

$$\int_c^d \int_a^\infty f(x, y)\, dx\, dy = \int_c^d F(y)\, dy = \int_a^\infty \int_c^d f(x, y)\, dy\, dx.$$

PRUEBA. Como f es continua sobre $\mathscr{R}$ y la integral impropia converge uniformemente sobre $[c, d]$, F es continua sobre $[c, d]$. Por tanto, $\int_c^d F(y)\, dy$ existe e $\int_c^d f(x, y)\, dy$ existe para todo $x \in [a, \infty\rangle$. Queremos

probar que

$$\int_c^d F(y)\,dy = \lim_{b\to\infty} \int_a^b \int_c^d f(x, y)\,dy\,dx\,;$$

es decir, queremos probar que para todo $\varepsilon > 0$ existe un número $N \geqslant a$ tal que

$$\left| \int_c^d F(y)\,dy - \int_a^b \int_c^d f(x, y)\,dy\,dx \right| < \varepsilon \quad \text{siempre que } b > N.$$

Ahora bien

$$\left| \int_c^d F(y)\,dy - \int_a^b \int_c^d f(x, y)\,dy\,dx \right| = \left| \int_c^d F(y)\,dy - \int_c^d \int_a^b f(x, y)\,dx\,dy \right|$$

$$= \left| \int_c^d \left[F(y) - \int_a^b f(x, y)\,dx \right] dy \right|.$$

Como $\int_a^\infty f(x, y)\,dx$ converge a $F(y)$ uniformemente sobre $[c, d]$, existe un numero $N \geqslant a$ tal que para todo $y \in [c, d]$

$$\left| F(y) - \int_a^b f(x, y)\,dx \right| < \frac{\varepsilon}{d-c} \quad \text{siempre que } b > N.$$

Por tanto, para todo $b > N$

$$\left| \int_c^d F(y)\,dy - \int_a^b \int_c^d f(x, y)\,dy\,dx \right|$$

$$= \left| \int_c^d \left[F(y) - \int_a^b f(x, y)\,dx \right] dy \right| < \frac{\varepsilon}{d-c}(d-c) = \varepsilon.$$

Lo que prueba que

$$\int_c^d F(y)\,dy = \lim_{b\to\infty} \int_a^b \int_c^d f(x, y)\,dy\,dx = \int_a^\infty \int_c^d f(x, y)\,dy\,dx.$$

5.7 Ejemplo. Sea $F(y) = \int_0^\infty e^{-ax} \operatorname{sen} yx\,dx$, $a > 0$. Integrando $\int_0^y F(u)\,du$, determínese $\int_0^\infty e^{-ax} \frac{1 - \cos yx}{x}\,dx$.

SOLUCIÓN. Tenemos

$$F(y) = \int_0^\infty e^{-ax} \operatorname{sen} yx\,dx = \lim_{b\to\infty} \left[\frac{-e^{-ax}}{a^2 + y^2}(a \operatorname{sen} yx + y\cos yx) \right]_0^b = \frac{y}{a^2 + y^2}.$$

Como $|e^{-ax}\operatorname{sen} yx| \leqslant e^{-ax}$ para todo $y \in \mathrm{R}$, según el criterio M de Weierstrass la convergencia es uniforme con respecto a y sobre R. Luego podemos cambiar el orden de integración:

$$\int_0^y F(u)\,du = \int_0^y \int_0^\infty e^{-ax}\operatorname{sen} ux\,dx\,du = \int_0^\infty e^{-ax}\int_0^y \operatorname{sen} ux\,du\,dx$$

$$= \int_0^\infty e^{-ax}\left[-\frac{\cos ux}{x}\right]_0^y dx = \int_0^\infty e^{-ax}\frac{1-\cos yx}{x}\,dx.$$

Por otra parte,

$$\int_0^y F(u)\,du = \int_0^y \frac{u}{a^2+u^2}\,du = \frac{1}{2}\ln(a^2+u^2)\Big]_0^y = \frac{1}{2}\ln\left(\frac{a^2+y^2}{a^2}\right).$$

De donde

$$\int_0^\infty e^{-ax}\frac{1-\cos yx}{x}\,dx = \frac{1}{2}\ln\left(\frac{a^2+y^2}{a^2}\right)$$

para $y \in \mathrm{R}$ y $a > 0$.

5.8 Teorema. *Si las funciones f y $D_2 f$ son continuas sobre la franja $\mathscr{R} = \{(x, y) \mid x \in [a, \infty\rangle, y \in [c, d]\}$, $\int_a^\infty f(x, y)\,dx$ converge a $F(y)$ sobre $[c, d]$, e $\int_a^\infty D_2 f(x, y)\,dy$ converge uniformemente sobre $[c, d]$, entonces F es diferenciable sobre $[c, d]$ y*

$$D_y \int_a^\infty f(x, y)\,dx = F'(y) = \int_a^\infty D_2 f(x, y)\,dx.$$

PRUEBA. Sea $g(y) = \int_a^\infty D_2 f(x, y)\,dx$ sobre $[c, d]$. Por el teorema 5.3 sabemos que g es continua sobre $[c, d]$ ya que $D_2 f$ es continua sobre $\mathscr{R}$ y la integral converge uniformemente sobre $[c, d]$. Por tanto, para $y \in [c, d]$, g es integrable sobre $[c, y]$ e

$$\int_c^y g(u)\,du = \int_c^y \int_a^\infty D_2 f(x, u)\,dx\,du.$$

Según el teorema 5.6, el orden de integración puede invertirse ya que $g(y) = \int_a^\infty D_2 f(x, y)\,dx$ converge uniformemente sobre $[c, y] \subset [c, d]$.

De donde

$$\int_c^y g(u)\,du = \int_a^\infty \int_c^y D_2 f(x, u)\,du\,dx = \int_a^\infty [f(x, y) - f(x, c)]\,dx$$

$$= F(y) - F(c).$$

Según el primer teorema fundamental del cálculo

$$F'(y) = g(y) = \int_a^\infty D_2 f(x, y)\,dx.$$

Problemas

1. Pruébese que las siguientes integrales convergen uniformemente sobre el intervalo especificado.

a) $\int_0^\infty e^{-x} \operatorname{sen} xy\,dx;\ y \in \mathrm{R}$ *b*) $\int_0^\infty y^x\,dx;\ y \in [0, a],\ a < 1$

c) $\int_0^\infty x^2 e^{xy}\,dx;\ y \leqslant -\delta < 0$ *d*) $\int_0^\infty \frac{e^{xy}}{2^x}\,dx;\ y \leqslant \ln a,\ a < 2$

e) $\int_0^\infty \frac{\operatorname{sen} xy}{x^2 + y^2}\,dx;\ y \geqslant \delta > 0$ *f*) $\int_0^\infty e^{-yx} \operatorname{sen} x\,dx;\ y \geqslant \delta > 0.$

2. Sea $F(a, b) = \int_0^\infty e^{-ax} \operatorname{sen} bx\,dx,\ a > 0,\ b \in \mathrm{R}$. Evalúese $F(a, b)$ y pruébese que se puede diferenciar con respecto tanto a a como a b y, por tanto, evaluar:

$$\int_0^\infty xe^{-ax} \operatorname{sen} bx\,dx$$

e

$$\int_0^\infty xe^{-ax} \cos bx\,dx.$$

3. Sea $F(y) = \int_0^\infty xy^{x-1}\,dx$. Pruébese que para una $a \in \langle 0, 1 \rangle$ fija y para $y \in \langle 0, 1 \rangle$, $\int_a^y F(u)\,du$ puede evaluarse cambiando el orden de integración. Evalúese $\int_a^y F(u)du$ y úsese el primer teorema fundamental del cálculo para determinar $F(y)$.

4. Sea $F(y) = \int_0^\infty e^{xy}\,dx$. Pruébese que F tiene derivadas de todos los órdenes sobre $\langle -\infty, 0\rangle$:

$$F^{(n)}(y) = \int_0^\infty x^n e^{xy}\,dx.$$

Evalúese $F(y)$ y úsese este valor para evaluar

$$\int_0^\infty x^n e^{xy}\,dx.$$

5. Sea $F(y) = \int_0^\infty \frac{\operatorname{sen} xy}{x} e^{-x}dx$. Encuéntrese $F'(y)$ y evalúese la integral para $F'(y)$. Partiendo del valor de $F'(y)$ determínese el valor de $F(y)$.

6. La función de Bessel modificada de segunda clase de orden cero puede definirse por

$$K_0(x) = \int_1^\infty \frac{e^{-xt}}{\sqrt{t^2-1}}\,dt \quad (x > 0).$$

Demuéstrese que K_0 satisface la ecuación diferencial (ecuación de Bessel modificada)

$$K_0''(x) + \frac{1}{x}K_0'(x) - K_0(x) = 0.$$

7. Para un entero positivo n, la función de Bessel modificada de segunda clase de orden n puede definirse por

$$K_n(x) = \frac{x^n}{(2n-1)!!}\int_0^\infty e^{-x\cosh t}\operatorname{senh}^{2n} t\,dt \quad (x > 0)$$

donde $(2n-1)!! = 1\cdot 3\cdot 5 \ldots (2n-1)$. Pruébese que K_n satisface la ecuación de Bessel modificada

$$K_n''(x) + \frac{1}{x}K_n'(x) - \left(1 + \frac{n^2}{x^2}\right)K_n(x) = 0.$$

Sugerencia: reemplácese $\cosh^2 t$ por $1 + \operatorname{senh}^2 t$, combínense las integrales, e intégrese una de las dos integrales resultantes por partes.

8. *a*) Pruébese que $\int_0^\infty \frac{dx}{(x^2+a)^{n+1}}$ es uniformemente convergente si $a \in \langle \delta, \infty\rangle$, donde $\delta > 0$ y n es un entero positivo cualquiera.

b) Pruébese que $\int_0^\infty \frac{dx}{x^2+a} = \frac{\pi}{2\sqrt{a}}$, $a \in \langle 0, \infty \rangle$.

c) Diferenciando respecto a a pruébese que

$$\int_0^\infty \frac{dx}{(x^2+a)^{n+1}} = \frac{(2n-1)!!\pi}{2^{n+1}n!} a^{-\frac{1}{2}(2n+1)}, \quad a \in \langle 0, \infty \rangle$$

para todo entero positivo n donde $(2n-1)!! = 1 \cdot 3 \cdot 5 \ldots (2n-1)$.

9. Evaluese $\int_0^\infty \frac{e^{-ax} - e^{-bx}}{x} dx$ evaluando $\int_0^\infty e^{-yx} dx$ y probando que es permisible integrar sobre el intervalo $[a, b]$, $a > 0$.

10. La función gamma (ejemplo 5.5) está definida por

$$\Gamma(y) = \int_0^\infty x^{y-1} e^{-x} dx \quad (y > 0).$$

Pruébese mediante integración por partes que

$$\Gamma(y+1) = y\Gamma(y).$$

Pruébese también que $\Gamma(1) = 1$. Dedúzcase de ello que para un n entero positivo

$$\Gamma(n+1) = n!.$$

6. EL VALOR DE UNA INTEGRAL CONVERGENTE

Si probamos que $\int_a^\infty f$ converge mediante el uso de la definición $\int_a^\infty f = \lim_{b \to \infty} \int_a^b f$, entonces obtenemos el valor de la integral infinita en el proceso de demostrar la convergencia. Este método puede usarse solamente en casos en que la integral $\int_a^b f$ puede evaluarse por uno de los métodos considerados en el cálculo elemental. Hemos visto que los teoremas de la sección previa pueden usarse también para la evaluación de las integrales impropias. Sin embargo, este método es de aplicabilidad muy limitada.

Si probamos que una integral impropia converge mediante la aplicación de uno de los criterios de la sección 3, entonces, en general, no conocemos el valor de la integral. Sabemos que podemos hallar una aproximación al valor de una integral convergente impropia $\int_a^\infty f$ con error menor que un

número positivo prefijado cualquiera mediante el cálculo de $\int_a^b f$ con tal de tomar b suficientemente grande. En esta sección daremos algunas estimaciones sobre cuán grande debe ser b para que $\int_a^b f$ se aproxime a $\int_a^\infty f$ con un grado especificado de precisión. Es entonces posible evaluar $\int_a^b f$ numéricamente y obtener así un valor aproximado de $\int_a^\infty f$.

Si probamos que $\int_a^\infty f$ es absolutamente convergente comparándola con la integral convergente $\int_a^\infty g$, entonces podemos obtener una estimación del error cometido al tomar $\int_a^b f$ como aproximación de $\int_a^\infty f$ como sigue: si $|f(x)| \leqslant g(x)$ siempre que $x > N$, entonces para $b_1 > b > N$

$$\left|\int_a^{b_1} f - \int_a^b f\right| = \left|\int_b^{b_1} f\right| \leqslant \int_b^{b_1} |f| \leqslant \int_b^{b_1} g.$$

Por tanto

6.1 $$\left|\int_a^\infty f - \int_a^b f\right| = \lim_{b_1 \to \infty} \left|\int_a^{b_1} f - \int_a^b f\right| \leqslant \int_b^\infty g$$

siempre que $b > N$. l numero $\left|\int_a^\infty f - \int_a^b f\right|$ se llama *error de truncación*. La desigualdad 6.1 nos da una cota superior del error de truncación.

6.2 Ejemplo. Estímese cuán grande debe ser b para que $\int_1^b \frac{\sqrt{x}}{3^x}\,dx$ se aproxime a la integral impropia $\int_1^\infty \frac{\sqrt{x}}{3^x}\,dx$ con error de truncación menor que 1×10^{-4}.

SOLUCIÓN. En el ejemplo 3.12, pág. 560, se probó que la integral impropia converge. La desigualdad

6.3 $$\frac{\sqrt{x}}{3^x} \leqslant \frac{1}{x^r}$$

se verifica para x suficientemente grande e $\int_b^\infty \frac{dx}{x^r}$ converge para $r > 1$.

Si tomamos $r = 2$, entonces 6.3 se verifica para todo $x > 0$. Para hacer

$$\int_b^\infty \frac{dx}{x^2} = \frac{1}{b} < 10^{-4}$$

es necesario tomar $b > 10^4$. Así pues, si usamos $r = 2$, parece que debemos tomaı b muy grande con el fin de asegurar una cota superior suficientemente pequeña del error de truncación. Si $r = 3$, entonces 6.3 se verifica para todo $x > 6$ e

$$\int_b^\infty \frac{dx}{x^3} = \frac{1}{2b^2} < 10^{-4}$$

para $b \geqslant 71$. Si $r = 4$, entonces 6.3 se verifica para $x > 9$ e

$$\int_b^\infty \frac{dx}{x^4} = \frac{1}{3b^3} < 10^{-4}$$

para $b \geqslant 15$. Si $r = 5$, entonces 6.3 se verifica para $x \geqslant 13$ mientràs que

$$\int_b^\infty \frac{dx}{x^5} = \frac{1}{4b^4} < 10^{-4}$$

para $b > 7.07$. Es, pues, suficiente tomar $b = 13$.

Si el integrando de una integral impropia cambia alternativamente de signo, entonces podemos construir una serie alternante partiendo de tal integral. El teorema 3.13, pág. 560, nos dice que si la serie alternante así construida satisface las condiciones del criterio de la serie alternante, entonces nuestra integral converge y su valor es igual a la suma de la serie alternante. Sabemos también que la suma de una serie alternante puede aproximarse por una suma parcial con un error de truncación menor que el valor absoluto del primer término omitido.

6.4 Ejemplo. Estímese cuán grande debe ser n para que $\displaystyle\int_{\pi/2}^{(n+3/2)\pi} \frac{\cos x}{x^2}\,dx$ se aproxime a $\displaystyle\int_{\pi/2}^{\infty} \frac{\cos x}{x^2}\,dx$ con error de truncación menor que ε.

SOLUCIÓN. Sea $\{b_k\} = (2k+1)\dfrac{\pi}{2}$ y

$$a_k = (-1)^{k+1} \int_{b_k}^{b_{k+1}} \frac{\cos x}{x^2}\,dx.$$

Entonces $\{a_k\}$ es una sucesión no creciente de términos positivos con

lím $a_k = 0$. De acuerdo con el teorema 3.13,

$$\int_{\pi/2}^{\infty} \frac{\cos x}{x^2}\,dx = \int_{\pi/2}^{3\pi/2} \frac{\cos x}{x^2}\,dx + \sum_{k=1}^{\infty} (-1)^{k+1} a_k$$

y el error de truncación es menor que

$$\left|\int_{(n+3/2)\pi}^{(n+5/2)\pi} \frac{\cos x}{x^2}\,dx\right| \leqslant \frac{1}{(n+\frac{3}{2})^2\pi^2}\left|\int_{(n+3/2)\pi}^{(n+5/2)\pi} \cos x\,dx\right|$$

$$= \frac{2}{(n+\frac{3}{2})^2\pi^2}.$$

El error de truncación es menor que ε si

$$n > \frac{1}{\pi}\sqrt{\frac{2}{\varepsilon}} - \frac{3}{2}.$$

En particular si $\varepsilon = 0.005$, entonces tomaríamos

$$n > \frac{20}{\pi} - \frac{3}{2} = 4.87.$$

Es decir, tomaríamos $n = 5$.

Nota. En el anterior ejemplo habríamos podido también probar la convergencia de la integral mediante el criterio de comparación:

$$\left|\frac{\cos x}{x^2}\right| \leqslant \frac{1}{x^2}.$$

Sin embargo, si hubiéramos estimado el error de truncación usando este criterio, habríamos llegado a la conclusión de que

$$\left|\int_{(n+3/2)\pi}^{\infty} \frac{\cos x}{x^2}\,dx\right| \leqslant \int_{(n+3/2)\pi}^{\infty} \frac{dx}{x^2} = \frac{1}{(n+\frac{3}{2})\pi} < \varepsilon$$

para $n > \dfrac{1}{\pi\varepsilon} - \dfrac{3}{2}$. En particular, para $\varepsilon = 0.0005$ habríamos tomado

$n > \dfrac{200}{\pi} - \dfrac{3}{2} = 62.66$ o $n = 63$.

Problemas

1. Úsese el criterio de comparación para estimar cuán grande debe tomarse b para que $\left|\int_a^\infty f - \int_a^b f\right|$ sea menor que ε en cada una de las siguientes integrales:

a) $\displaystyle\int_1^\infty \frac{dx}{x\sqrt{1+x^2}}$ *b)* $\displaystyle\int_0^\infty \frac{ue^u\,du}{(1+e^{2u})^2}$

c) $\displaystyle\int_1^\infty \frac{\ln x}{1+x^2}\,dx$ *d)* $\displaystyle\int_1^\infty \frac{dy}{y\sqrt{3y^2+2y+1}}$

e) $\displaystyle\int_1^\infty \frac{\ln x\,dx}{(1+x^2)^2}$ *f)* $\displaystyle\int_0^\infty \frac{dx}{1+\operatorname{sen} x+e^x}$

g) $\displaystyle\int_0^\infty e^{-x^2}\,dx$ *h)* $\displaystyle\int_1^\infty e^{-x}\ln x\,dx$.

2. Úsese el criterio de comparación para estimar cuán próximo a cero debe escogerse δ para que $\left|\int_a^b f - \int_{a+\delta}^b f\right|$ o $\left|\int_a^b f - \int_a^{b-\delta} f\right|$ sea menor que ε en cada una de las siguientes integrales:

a) $\displaystyle\int_0^1 \frac{\operatorname{sen} x}{x^{3/2}}\,dx$ *b)* $\displaystyle\int_0^{\pi/2} \frac{\sqrt{x}}{x+\operatorname{sen} x}\,dx$

c) $\displaystyle\int_0^1 \frac{dx}{(1+x)\sqrt{x}}$ *d)* $\displaystyle\int_0^1 \frac{dx}{\sqrt{1+x^2}}$.

3. Estímese cuán grande debe escogerse b para que $\left|\int_a^\infty f - \int_a^b f\right|$ sea menor que ε mediante la consideración de la serie alternante asociada:

a) $\displaystyle\int_1^\infty \frac{\operatorname{sen} x}{x^2}\,dx$ *b)* $\displaystyle\int_0^\infty \frac{\operatorname{sen} x}{1+x^2}\,dx$

c) $\displaystyle\int_{\pi/2}^\infty \frac{\operatorname{sen} x}{\ln x}\,dx$ *d)* $\displaystyle\int_0^\infty e^{-x}\operatorname{sen} x\,dx$.

7. RESUMEN

En este capítulo hemos considerado ciertas generalizaciones de la noción de integral. En el capítulo 6, la integral definida $\int_a^b f$ se generalizó primero reemplazando el intervalo $[a, b]$ en R por un intervalo en R^n y después reemplazando $[a, b]$ por ciertos conjuntos acotados más generales en R^n. Aquí las generalizaciones tomaron direcciones diferentes.

Hemos estudiado integrales impropias —integrales impropias de primera clase en que el intervalo de integración es infinito e integrales impropias de segunda clase en que el intervalo es finito, pero el integrando no es acotado. Vimos que en muchos respectos las integrales impropias son completamente análogas a las series infinitas. Derivamos varios criterios de convergencia para las derivadas impropias y se consideraron algunos métodos para su evaluación.

Como otra generalización de la integral, estudiamos también las integrales, tanto propias como impropias, dependientes de un parámetro:

$$F(y) = \int_a^b f(x, y)\, dx$$

donde a y b pueden ser números reales, funciones de y o más o menos infinito. Las integrales dependientes de un parámetro ya se nos habían presentado en conexión con las integrales iteradas. Aquí hemos dado condiciones suficientes de continuidad, diferenciabilidad e integrabilidad de F.

Problemas de repaso

1. Evalúense las siguientes integrales impropias.

a) $\displaystyle\int_0^{\pi/2} \frac{\operatorname{sen} x}{1 - \cos x}\, dx$ *b*) $\displaystyle\int_0^{\infty} e^{-x} \operatorname{sen} x\, dx$

c) $\displaystyle\int_0^{\infty} e^{-2x}\, dx$ *d*) $\displaystyle\int_2^{5} \frac{dx}{x\sqrt{x^2-4}}$

e) $\displaystyle\int_0^{\infty} \frac{dx}{a^2+b^2x^2}$ *f*) $\displaystyle\int_0^{\infty} \frac{dx}{(1+x)^{3/2}}$.

2. Pruébese que la región limitada superiormente por la hipérbola $xy = 1$, inferiormente por el eje X, a la izquierda por la recta $x = 1$, y no limitada hacia la derecha, no tiene área alguna definible. Pruébese también que si esta región se hace girar alrededor del eje X, entonces el sólido de revolución generado tiene un volumen definible. Encuéntrese este volumen.

3. Encuéntrese el área total limitada por la gráfica de la ecuación

$$x^2y^2+2x^2-4y^2 = 0$$

y sus asíntotas.

4. Para m y n positivos, puede definirse la función beta por

$$B(m, n) = \int_0^\infty \frac{x^{m-1}\,dx}{(1+x)^{m+n}}.$$

a) Pruébese mediante integración por partes que

$$B(m, n) = \frac{m-1}{m-1+n} B(m-1, n) \quad \text{para } m > 1, \ n > 0.$$

b) Dedúzcase que para m y n enteros positivos,

$$B(m, n) = \frac{(m-1)!\,(n-1)!}{(m+n-1)!} = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)}$$

donde Γ es la función gamma que introdujimos en el ejemplo 5.5, pág. 570.

5. Sean N y D polinomios de grados n y d respectivamente. Demuéstrese que si a es mayor que el mayor cero (real) de D, entonces $\int_a^\infty \frac{N(x)}{D(x)}dx$ converge si y sólo si $d > n+1$.

6. Determínese si las siguientes integrales convergen o divergen.

a) $\int_0^\infty \frac{\arctan x}{1+x^2}\,dx$ *b*) $\int_0^\infty \frac{dx}{\sqrt{1+x^3}}$

c) $\int_0^\infty \frac{x^3\,dx}{\sqrt{x^7+1}}$ *d*) $\int_0^{\pi/2} \frac{\cos x}{\sqrt{x}}\,dx$

e) $\int_0^{\pi/2} \frac{x^2}{\operatorname{sen}^3 x}\,dx$ *f*) $\int_0^1 \frac{\operatorname{sen} x}{x\sqrt{1-x^2}}\,dx.$

7. Diferénciense cada una de las siguientes funciones:

a) $\int_0^\pi \cos(xy^2)\,dx$ *b*) $\int_3^7 \operatorname{sen}(x-y)\,dx$

c) $\int_0^{\pi/y^2} \cos(xy^2)\,dx$ *d*) $\int_x^{x^2} e^{-xy}(y-x)^4\,dy.$

8. Sea $F(y) = \int_0^y \operatorname{sen}(xy)\,dx$. Evalúese F y, luego, diferénciese tanto en la forma integral como en la forma evaluada. Partiendo de ello, obténgase el valor de

$$\int_0^y x\cos(xy)\,dx.$$

9. Sea $F_n(y) = \int_0^y (y-x)^n \cos x\,dx$ donde n es un entero positivo.

b) Encuéntrense $F_n'(y)$, $F_n''(y)$, ... sin integración, y pruébese que

$$F_n^{(n)}(y) = n!\operatorname{sen} y.$$

b) Basándose en los resultados de la parte *a*, pruébese que

$$F_n(y) = n!f(y) + P_{n-1}(y)$$

donde $P_{n-1}(y) = a_0 + a_1 y + \ldots + a_{n-1}y^{n-1}$ es un polinomio de grado, cuando más, $n-1$ y

$$f(y) = \begin{cases} (-1)^{n/2}\operatorname{sen} y & \text{para } n \text{ par} \\ (-1)^{\frac{1}{2}(n+1)}\cos y & \text{para } n \text{ impar.}\end{cases}$$

c) Obsérvese que $F_n(0) = F_n'(0) = \ldots = F_n^{(n)}(0) = 0$. Usese este hecho para determinar los coeficientes de $P_5(y)$ y luego evalúese

$$F_6(y) = \int_0^y (y-x)^6 \cos x\,dx.$$

10. Pruébese que las siguientes integrales convergen uniformemente sobre el intervalo que, en cada caso, se señala.

a) $\int_1^\infty \frac{\cos x}{x^2+y^2}\,dx$; $y \in \mathrm{R}$ *b*) $\int_0^\infty \frac{\operatorname{sen}^2(yx)}{yx^2}\,dx$; $y \geqslant \delta > 0$

c) $\int_0^\infty te^{-st}\,dt$; $s \geqslant \delta > 0$ *d*) $\int_0^\infty t^n e^{-st}\,dt$; $n \geqslant 0$, $s \geqslant \delta > 0$.

11. Sea $\bar{f}(s) = \int_0^\infty e^{-st}\,dt$, para $s \in \langle 0, \infty\rangle$.

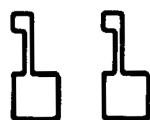
a) Pruébese que $\bar{f}(s) = \frac{1}{s}$ para $s \in \langle 0, \infty\rangle$.

b) Pruébese que es permisible diferenciar $\bar{f}(s)$ n veces bajo el signo integral para $s \in \langle 0, \infty\rangle$.

c) Dedúzcase que para un entero positivo cualquiera n y $s \in \langle 0, \infty\rangle$,

$$\int_0^\infty t^n e^{-st}\,dt = \frac{n!}{s^{n+1}}.$$





Ecuaciones diferenciales

1. INTRODUCCIÓN

Este capítulo es una breve introducción a las ecuaciones diferenciales y en él queremos conocer lo que una ecuación diferencial es, y comprender lo que quiere decirse cuando se habla de una solución de una ecuación diferencial. Nos limitamos a algunos tipos sencillos de ecuaciones que pueden resolverse en términos de funciones elementales o cuya solución puede expresarse analíticamente en términos de una integral definida. Los problemas y los ejemplos nos demuestran algunas de las formas en que se aplican las ecuaciones diferenciales a problemas concretos, y veremos cómo las ecuaciones diferenciales, más ciertas condiciones iniciales, determinan soluciones únicas. Las ecuaciones diferenciales que hemos elegido como

objeto de estudio son también las que sirven como ejemplos elementales importantes en ciencia e ingeniería.

El problema central en las ecuaciones diferenciales puede describirse en toda su generalidad como el estudio de un conjunto de funciones definido por el requerimiento de que la función y algunas de sus derivadas tengan propiedades especiales. Por ejemplo, el conjunto puede ser el conjunto de todas las funciones reales con la propiedad de que la derivada de cada una de las funciones sobre $\langle -\infty, \infty \rangle$ sea la propia función. Este requerimiento puede expresarse por la ecuación

1.1 $$y' = y \quad \text{sobre } \langle -\infty, \infty \rangle$$

o por la ecuación escrita como regla de correspondencia

1.1′ $$y'(x) = y(x), \quad x \in \langle -\infty, \infty \rangle .$$

La ecuación 1.1 se llama "ecuación diferencial", y cualquier función con esta propiedad se llama "solución" de la ecuación diferencial. Sabemos que $D(\exp) = \exp$. Por tanto, $y = \exp$ es una solución de la ecuación 1.1. Podemos también decir que $y(x) = e^x$ es una solución, y entender por esto que estamos enunciando la regla de correspondencia de una solución.

Sea c una constante cualquiera y $y(x) = ce^x$, entonces $y'(x) = ce^x = y(x)$. Por tanto, $y(x) = ce^x$ es una solución para cualquier constante c. ¿Tiene esta ecuación diferencial algunas otras soluciones? Para contestar a esta pregunta supongamos que u es una solución de la ecuación 1.1. Entonces

$$u'(x) - u(x) = 0$$

y

$$e^{-x}(u'(x) - u(x)) = D_x(e^{-x}u(x)) = 0 .$$

De donde

$$e^{-x}u(x) = c \quad \text{y} \quad u(x) = ce^x .$$

Por tanto, toda solución de la ecuación 1.1 sobre $\langle -\infty, \infty \rangle$ es de la forma ce^x, donde c es una constante y ce^x se llama "solución general" de la ecuación 1.1. Toda solución es de la forma de la solución general, y todas las funciones de esta forma son soluciones.

El conjunto de funciones definido por una ecuación diferencial puede restringirse más por las que se llaman "condiciones iniciales" o "condiciones de frontera". Volviendo a nuestro ejemplo, consideremos

1.2 $$y' = y \quad \text{sobre } \langle -\infty, \infty \rangle, \ y(0) = 1 .$$

Deseamos encontrar todas las soluciones de la ecuación 1.1 que satisfacen la condición inicial $y(0) = 1$. Sabemos que la solución general de la ecuación 1.1 es $y(x) = ce^x$. Como $y(0) = c = 1$, $y(x) = e^x$ es la única solución de la ecuación 1.2.

Otros ejemplos de ecuaciones diferenciales son:
La ecuación que describe el movimiento de un péndulo:

1.3 $$\ddot{\theta} + \text{sen} \circ \theta = 0 \quad (\ddot{\theta}(t) + \text{sen}\, \theta(t) = 0).$$

La ecuación diferencial para un circuito eléctrico que consiste en una inductancia (L), una resistencia (R) y una capacitancia (C) dispuestas en serie $\left(q = \text{carga};\ \dot{q} = \dfrac{dq}{dt} = \text{corriente};\ E, \text{ voltaje aplicado}\right)$:

1.4 $$L\ddot{q} + R\dot{q} + \frac{1}{C}q = E \quad \left(L\ddot{q}(t) + R\dot{q}(t) + \frac{1}{C}q(t) = E(t)\right).$$

Un par de ecuaciones diferenciales en dos funciones incógnitas x y y:

1.5 $$\begin{array}{l} \dot{x} = ax + by \\ \dot{y} = cx + dy \end{array} \quad \left(\begin{array}{l} \dot{x}(t) = ax(t) + by(t) \\ \dot{y}(t) = cx(t) + dy(t) \end{array}\right).$$

La ecuación de Newton (1671) para el movimiento de una partícula en el campo gravitatorio terrestre (el problema de los dos cuerpos):

1.6 $$m\ddot{\mathbf{r}} = -k\,\frac{\mathbf{r}}{r^3}, \quad \text{donde } r = |\mathbf{r}|.$$

La ecuación de Bessel. (El primer estudio sistemático de las soluciones fue dado por Bessel en 1824):

1.7 $$x^2 y''(x) + xy'(x) + (x^2 - n^2)\, y(x) = 0.$$

Nota. Esta ecuación se escribe a menudo en la forma $x^2 \dfrac{d^2 y}{dx^2} + x\dfrac{dy}{dx} + (x^2 - n^2)y = 0$ o, simplemente, $x^2 y'' + xy' + (x^2 - n^2)y = 0$. Aunque la notación es incompleta, no debe haber dentro del contexto de las ecuaciones diferenciales ningún mal entendido. En lugar de la ecuación 1.3 podemos escribir $\dfrac{d^2\theta}{dt^2} + \text{sen}\,\theta = 0$ o $\ddot{\theta} + \text{sen}\,\theta = 0$. Se entiende, entonces, que θ es una función y que sen θ es una composición de funciones. Si quisiéramos que fuera el producto de funciones, escribiríamos $\ddot{\theta} + (\text{sen}\, t)\theta = 0$. Además, si en la ecuación 1.5 no supusiéramos que a, b, c y d eran constantes escribiríamos $\dot{x} = a(t)x + b(t)y$, $\dot{y} = c(t)x + d(t)y$.

La ecuación de Laplace (1787), introducida por Laplace en una memoria sobre los anillos de Saturno:

1.8 $$D_1{}^2 u + D_2{}^2 u + D_3{}^2 u = 0 \quad \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0\right).$$

La eacución de Mathieu (1868), que aparece en el estudio de las vibraciones de una membrana elíptica:

1.9 $$\frac{dy}{dx} + (a+b\cos 2x)y = 0.$$

La ecuación de Van der Pol (1922), la ecuación diferencial de un oscilador triódico:

1.10 $$\ddot{x}+\varepsilon(x^2-1)\dot{x}+x = 0.$$

Una ecuación diferencial aproximada para la vibración torsional de una estructura mecánica sujeta a amortiguamiento aerodinámico y friccional (un estudio de esta ecuación ayuda a comprender la sensacional falla del puente colgante de Tacoma en 1940):

1.11 $$\ddot{\theta}+(f(\dot{\theta})+g(\theta))\dot{\theta}+\theta = 0.$$

Las ecuaciones diferenciales para un sistema automático de control:

1.12 $$\begin{aligned}\ddot{\varepsilon}+a\dot{\varepsilon}+b\varepsilon &= cz\\ \dot{z} &= f(k_1 y-k_2 z)\\ \ddot{y}+my &= d\dot{\varepsilon}+h\varepsilon.\end{aligned}$$

Las ecuaciones diferenciales se dividen en dos clases: 1) ecuaciones diferenciales ordinarias, que definen funciones de una sola variable real, y 2) ecuaciones diferenciales parciales, como la ecuación 1.8, que definen funciones de dos o más variables reales. Con excepción de la ecuación 1.8, todas las anteriores ecuaciones son ecuaciones diferenciales ordinarias.

Nuestro interés primario en este capítulo es el estudio de las ecuaciones diferenciales ordinarias, y serán de la forma

1.13 $$\mathbf{y}^{(m)} = \mathbf{F}\circ(I, \mathbf{y}, \mathbf{y}', \cdots, \mathbf{y}^{(m-1)}),$$
$$(\mathbf{y}^{(m)}(t) = \mathbf{F}(t, \mathbf{y}(t), \mathbf{y}'(t), \cdots, \mathbf{y}^{(m-1)}(t))$$

donde $\mathbf{y}^{(k)} = D^k\mathbf{y}$ y $\mathbf{F}$ es una función de R^{nm+1} en R^n. Una función $\mathbf{g}$ de R en R^n se dice que es una *solución de la ecuación* 1.13 sobre un intervalo $\mathscr{I}$ si

$$\mathbf{g}^{(m)} = \mathbf{F}\circ(I, \mathbf{g}, \mathbf{g}', \cdots, \mathbf{g}^{(m-1)}) \quad \text{sobre } \mathscr{I};$$

es decir,

$$\mathbf{g}^{(m)}(t) = \mathbf{F}(t, \mathbf{g}(t), \mathbf{g}'(t), \ldots, \mathbf{g}^{(m-1)}(t)) \quad \text{para todo } t\in\mathscr{I}.$$

El rango de $\mathbf{F}$ está en R^n y la ecuación 1.13 representa un sistema de n ecuaciones. El número m es el orden de la derivada de orden más alto en la ecuación, y la ecuación 1.13 se dice que es *un sistema de n ecuaciones diferenciales ordinarias de orden m.*

Problemas

Verifíquese que cada una de las siguientes funciones es una solución sobre $\mathscr{I}$ de la ecuación diferencial dada.

1. $y = \text{sen}$, $y''+y = 0$, $\mathscr{I} = \langle -\infty, \infty \rangle$

2. $y = c_1 \text{ sen} + c_2 \cos$, $y''+y = 0$, $\mathscr{I} = \langle -\infty, \infty \rangle$

3. $y(x) = A \text{ sen } (x+\delta)$, $y''+y = 0$, $\mathscr{I} = \langle -\infty, \infty \rangle$

4. $y(t) = c_1 \text{ sen } t + c_2 \cos t + \frac{1}{2}e^t$, $y''+y = e^t$, $\mathscr{I} = \langle -\infty, \infty \rangle$

5. $y(x) = \text{senh } x = \dfrac{e^x - e^{-x}}{2}$, $y''-y = 0$, $\mathscr{I} = \langle -\infty, \infty \rangle$

6. $y(x) = c_1 e^x + c_2 e^{-x}$, $y''-y = 0$, $\mathscr{I} = \langle -\infty, \infty \rangle$

7. $y(x) = \dfrac{1}{\sqrt{1-x}}$, $2y'-y^3 = 0$, $\mathscr{I} = \langle -\infty, 1 \rangle$

8. $y(x) = \dfrac{1}{x^2}$, $x^2 y''+3xy' = 0$, $\mathscr{I} = \langle 0, \infty \rangle$

9. $\mathbf{r}(t) = (\cos \omega t, \text{ sen } \omega t)$, $\ddot{\mathbf{r}} + \omega^2 \mathbf{r} = \mathbf{0}$, $\mathscr{I} = \langle -\infty, \infty \rangle$

10. $z(t) = e^{\beta t}(\cos \omega t + i \text{ sen } \omega t)$, $\dot{z} = (\beta + i\omega)z$, $\mathscr{I} = \langle -\infty, \infty \rangle$

11. Pruébese que $e^{\lambda_1 t}$ es una solución de

$$a\ddot{x}+b\dot{x}+cx = 0 \quad (b^2-4ac \geqslant 0)$$

si y sólo si λ_1 es una raíz de $a\lambda^2+b\lambda+c = 0$.

12. Verifíquese que

$$u(\mathbf{r}) = u(x, y, z) = \frac{1}{\sqrt{x^2+y^2+z^2}} = \frac{1}{|\mathbf{r}|}, \quad \mathbf{r} = (x, y, z) \neq \mathbf{0}$$

es una solución de la ecuación de Laplace

$$D_1{}^2 u + D_2{}^2 u + D_3{}^2 u = 0.$$

13. Verifíquese que

$$u(\mathbf{r}) = \frac{1}{|\mathbf{r}|}, \quad \mathbf{r} \neq \mathbf{0},$$

es una solución de

$$\nabla u = -\frac{\mathbf{r}}{|\mathbf{r}|^3}$$

$$(\nabla u = \text{grad } u = (D_1 u, D_2 u, D_3 u)).$$

Supongamos que f es continua sobre un intervalo $\mathscr{I}$ y que $0 \in \mathscr{I}$. Verifíquese que:

14. $x(t) = \int_0^t (t-s) f(s) ds$ es una solución sobre $\mathscr{I}$ de $\ddot{x} = f$ que satisface $x(0) = \dot{x}(0) = 0$.

15. $x(t) = \int_0^t \frac{(t-s)^{n-1}}{(n-1)!} f(s) ds$ es una solución sobre $\mathscr{I}$ de $x^{(n)} = f$ que satisface $x(0) = \dot{x}(0) = \cdots = x^{(n-1)}(0) = 0$.

16. $x(t) = \int_0^t \operatorname{sen}(t-s) f(s) ds$ es una solución sobre $\mathscr{I}$ de $\ddot{x} + x = f$ que satisface $x(0) = \dot{x}(0) = 0$.

17. (Problema 3, pág. 567.) Si A es una solución sobre $\mathscr{I}$ de $\dot{x} + bx = 0$ que satisface $A(0) = 1$, entonces

$$x(t) = x_0 A(t) + \int_0 A(t-s) f(s) ds$$

es una solución sobre $\mathscr{I}$ de $\dot{x} + bx = f$ que satisface $x(0) = x_0$.

2. LA ECUACIÓN $y' = f$

La historia del estudio de las ecuaciones diferenciales comienza en la última parte del siglo diecisiete con la fundación del cálculo por Isaac Newton y Gottfried Leibniz.[1] En 1671 Leibniz introdujo la terminología "aequatic diferencialis". Newton y Leibniz habían descubierto independientemente los teoremas fundamentales del cálculo, y estos teoremas (capítulo 3, pág. 132) proporcionaron el teorema básico de existencia y unicidad de las ecuaciones diferenciales.

[1] Es cierto, sin embargo, que John Napier (1550-1617) definió la función logarítmica cinemáticamente, y su definición era equivalente a definir la función L —logaritmo de Napier— como la solución $y(x) = L(x)$ de la ecuación diferencial

$$y'(x) = -\frac{10^7}{x}, \quad y(10^7) = 0.$$

Esto es equivalente a definir $L(x) = -\int_{10^7}^x \frac{10^7}{t} dt$, y, por tanto,

$$L(x) = -10^7 \ln \frac{x}{10^7} = 10^7 \ln \frac{10^7}{x}.$$

Napier calculó una tabla de logaritmos mediante un método de aproximación derivado de esta descripción cinemática de su definición de L.

2.1 Teorema. *Si f es continua sobre un intervalo $\mathscr{I}$, si $x_0 \in \mathscr{I}$, y si y_0 es un número cualquiera en $\langle -\infty, \infty \rangle$, entonces la ecuación diferencial*

2.2 $$y' = f$$

tiene una solución única sobre $\mathscr{I}$ que satisface $y(x_0) = y_0$. Esta solución es

$$y(x) = y_0 + \int_{x_0}^{x} f.$$

PRUEBA. Supongamos que y es una solución de la ecuación 2.2 que satisface $y(x_0) = y_0$. Entonces, de acuerdo con el segundo teorema fundamental

$$\int_{x_0}^{x} f = \int_{x_0}^{x} y' = y(x) - y(x_0) = y(x) - y_0$$

y

2.3 $$y(x) = y_0 + \int_{x_0}^{x} f.$$

El primer teorema fundamental nos dice que la función definida por 2.3 es una solución y, claramente, esta función satisface la condición inicial. Y esto completa la prueba.

Nota. La notación de la integral indefinida se usa con frecuencia en las ecuaciones diferenciales. Si f es continua sobre $\mathscr{I}$, entonces $F = \int f$ es equivalente a decir que F es una solución de $y' = f$ sobre $\mathscr{I}$. Así pues, en este caso, $\int f$ puede usarse para denotar una solución y es meramente una notación que denota una solución. Podemos, por ejemplo, decir que $y = \int f + c$ es la solución general de $y' = f$. Esto significa que si conocemos una solución sobre $\mathscr{I}$, entonces cualquier constante más esa solución, es una solución sobre $\mathscr{I}$ y todas las soluciones son de esa forma.

2.4 Ejemplo. Se dispara verticalmente un proyectil desde la superficie de la tierra con una velocidad inicial de 1 000 pies por segundo. Prescindiendo del efecto de la atmósfera y suponiendo la fuerza de la gravedad constante, estímese la máxima altura alcanzada por el proyectil.

SOLUCIÓN. Sea $y(t)$ la distancia (en pies) del proyectil sobre la tierra t segundos después de haberla dejado. Aquí las condiciones iniciales son $y(0) = 0$ y $\dot{y}(0) = 1\,000$. Sea g la aceleración (pies/segundo^2) de la gravedad. Entonces

$$\ddot{y} = -g,$$

$$\int_0^t \ddot{y} = \dot{y}(t) - \dot{y}(0) = -\int_0^t g = -gt.$$

Luego

$$\dot{y}(t) = \dot{y}(0) - gt$$

y

$$y(t) = y(0) + \int_0^t (\dot{y}(0) - g\tau)\, d\tau$$
$$= y(0) + \dot{y}(0)t - \tfrac{1}{2}gt^2.$$

Por tanto

$$y(t) = y(0) + \dot{y}(0)t - \tfrac{1}{2}gt^2 = \dot{y}(0)t - \tfrac{1}{2}gt^2.$$

En $t = \dfrac{\dot{y}(0)}{g}$, $\dot{y}(t) = 0$, y es claro [ya que $\ddot{y}(t) < 0$ para todo t] que

$$y_{\text{máx}} = \dot{y}(0)\frac{\dot{y}(0)}{g} - \frac{\dot{y}^2(0)}{2g} = \frac{\dot{y}^2(0)}{2g}.$$

La constante g es aproximadamente 32.2 pies/segundo^2 y $\dot{y}(0) = 1\,000$ pies/segundo. Luego

$$y_{\text{máx}} = 155 \times 10^2 \text{ pies} = 15\ 500 \text{ pies}.$$

2.5 Ejemplo. Resuélvase

$$y' = \frac{1}{y}.$$

SOLUCIÓN. Supongamos que la ecuación diferencial tiene una solución sobre algún intervalo. Entonces, como $yy' = 1$, tenemos

$$yy' = \tfrac{1}{2}D(y^2) = 1,$$

y

$$y^2(x) = 2x + c.$$

De donde (como y debe ser diferenciable)

$$y(x) = \sqrt{2x+c} \qquad \text{para} \quad x > -\frac{c}{2}$$

o bien

$$y(x) = -\sqrt{2x+c} \quad \text{para} \quad x > -\frac{c}{2}.$$

Es fácil verificar ahora que estas son soluciones sobre $\langle -c/2, \infty\rangle$, y que, por tanto, la solución general sobre $\left\langle -\dfrac{c}{2}, \infty \right\rangle$ es

$$y(x) = \sqrt{2x+c}$$

o bien

$$y(x) = -\sqrt{2x+c}\cdot$$

Así pues, si, por ejemplo, $y(0) = 1$, entonces $y(x) = \sqrt{2x+1}$ sobre $\langle -\frac{1}{2}, \infty\rangle$.

Problemas

1. Determínense todas las soluciones sobre $\langle -\infty, \infty\rangle$ de

a) $y' = 1 + \cos$ *b*) $y'(x) = x^2 - 3x^5$
c) $u'(t) = e^{-t^2}$ *d*) $x'(s) = \ln(1+s^2)$.

2. Determínese la función definida por

a) $y'(x) = 1/x$ sobre $\langle 0, \infty\rangle$, $y(1) = 0$
b) $y'(x) = 1/x$ sobre $\langle -\infty, 0\rangle$, $y(-1) = 0$

c) $y' = f$, donde $f(x) = \dfrac{\operatorname{sen} x}{x}$ cuando $x \neq 0$ y $f(0) = 1$, y $y(0) = 10$.

3. La función sgn (léase "signo") está definida por

$$\operatorname{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases}$$

Considérese la ecuación diferencial

$$y'(x) = \operatorname{sgn} x.$$

a) Determínense todas las soluciones sobre $\langle 0, \infty\rangle$.
b) Determínense todas las soluciones sobre $\langle -\infty, 0\rangle$.
c) Determínense todas las soluciones sobre $\langle -\infty, \infty\rangle$.

4. Si se supone, además, que satisface la ecuación diferencial del problema 3 sobre todo intervalo donde sgn es continuo, que y es continua, determínese y dado que $y(0) = y_0$.

5. Pruébese que $e^{ax}[y'(x)+ay(x)] = D_x(e^{ax}y(x))$. Partiendo de ello resuélvase (es decir, determínense todas las soluciones) la ecuación diferencial

a) $y' - 2y = 0$ *b*) $y' = 2y+6$ *c*) $y'(x)+3y(x) = e^x$

d) $\begin{cases} y'(x)+4y(x) = x, \\ y(0) = 1 \end{cases}$ *e*) $\begin{cases} y'+y = \operatorname{sen}, \\ y(0) = 0 \end{cases}$

f) $\ddot{x}+\dot{x} = 0$
g) $\ddot{x}(t)+3\dot{x}(t) = 30$, $x(0) = 1$, $\dot{x}(0) = 1$.

6. Resuélvase

$$\dot{\mathbf{r}}(t)+a\mathbf{r}(t) = \mathbf{0}$$

dado que $\mathbf{r}(0) = \mathbf{c}$.

7. Resuélvanse

a) $y''(t) = 32$ $\quad y(0) = 19$ $\quad y'(0) = 28$
b) $y''(t) = t+4$ $\quad y(0) = 2$ $\quad y'(0) = 0$
c) $y''(t) = \text{sen } t$ $\quad y(0) = 0$ $\quad y'(0) = 1.$

8. Supóngase que f es continua sobre un intervalo $\mathscr{I}$ y que $0 \in \mathscr{I}$. Intercambiando el orden de integración pruébese que

$$\int_0^x \int_0^t f(s)\, ds\, dt = \int_0^x (x-s)\, f(s)\, ds\,.$$

Pruébese, partiendo de ello, que *la* solución de

$$y'' = f$$

que satisface $y(0) = y_0$ y $y'(0) = \dot{y}_0$ es

$$y(x) = \dot{y}_0 x + y_0 + \int_0^x (x-s)\, f(s)\, ds\,.$$

9. Resuélvanse

a) $y(x)\, y'(x) = x$
b) $y(x)\, y'(x) = x \quad y(0) = 1$
c) $y(x)\, y'(x) = -x \quad y(0) = 1$
d) $y'(x) = \dfrac{y(x)}{x}$ sobre $\langle 0, \infty \rangle$.

10. Interprétense las ecuaciones diferenciales de los problemas $9b$ y $9d$ geométricamente como condiciones sobre la gráfica de y.

11. ¿Hay funciones g continuas sobre $\langle -\infty, \infty \rangle$ que satisfagan sobre $\langle -\infty, \infty \rangle$

a) $x^2 = \int_a^x g$? $\qquad$ *b*) $5e^x = 1 + \int_a^x g$? $\qquad$ *c*) $e^{2x} = \int_a^{x'} g$?

12. Verifíquese que $y(x) = 0$ y $y(x) = \frac{1}{9}x^3$ son soluciones sobre $\langle -\infty, \infty \rangle$ de $y'(x) = \sqrt{xy(x)}$ que satisfacen $y(0) = 0$.

3. LA ECUACIÓN DIFERENCIAL LINEAL DE PRIMER ORDEN $x' + px = q$

La *ecuación diferencial lineal general de primer orden* es una ecuación de la forma

3.1 $$Ax' + Bx = C \quad (A(t)\, x'(t) + B(t)\, x(t) = C(t))$$

donde A, B y C son funciones. Supondremos que A, B y C son continuas sobre un intervalo $\mathscr{I}$ y que A no tiene ceros sobre $\mathscr{I}$. En este caso, la ecuación 3.1 es equivalente sobre $\mathscr{I}$ a (es decir, tiene las mismas soluciones sobre $\mathscr{I}$ que)

$$x' + \frac{B}{A}x = \frac{C}{A},$$

que es una ecuación de la forma

3.2 $$x' + px = q \quad (x'(t) + p(t)\,x(t) = q(t))$$

donde p y q se supone que son continuas sobre un intervalo $\mathscr{I}$. Probaremos ahora que multiplicando ambos lados de esta ecuación por una función, a la que llamamos "factor de integración", el primer miembro de la ecuación se convierte en una derivada y la ecuación se reduce a una de la forma

$$y' = g.$$

Para las ecuaciones lineales de primer orden se encuentra fácilmente un factor de integración. Sea $t_0 \in \mathscr{I}$ y

$$P = \int_{t_0} p;$$

es decir,

$$P(t) = \int_{t_0}^{t} p(s)\,ds \quad \text{para todo } t \in \mathscr{I}.$$

Entonces

$$P' = p \quad \text{sobre } \mathscr{I},$$

y si x es una ecuación diferenciable cualquiera sobre $\mathscr{I}$,

3.3 $$e^{P(t)}(x'(t) + p(t)\,x(t)) = D_t(e^{P(t)}\,x(t)).$$

Como $e^{P(t)}$ nunca es cero, multiplicando ambos lados de 3.2 por el factor de integración $e^{P(t)}$, encontramos que la ecuación 3.2 es equivalente a

$$D_t(e^{P(t)}\,x(t)) = e^{P(t)}\,q(t).$$

Si $x(t_0) = x_0$ —y nótese que, por definición, $P(t_0) = 0$— entonces esta ecuación tiene, de acuerdo con el teorema 2.1, la solución única sobre $\mathscr{I}$:

$$e^{P(t)}\,x(t) = e^{P(t_0)}x_0 + \int_{t_0}^{t} e^{P(s)}\,q(s)\,ds$$

o bien

3.4 $$x(t) = e^{-P(t)}x_0 + e^{-P(t)}\int_{t_0}^{t} e^{P(s)}\,q(s)\,ds.$$

Así pues, la multiplicación por el factor de integración $e^{P(t)}$, donde $P(t) = \int_{t_0}^{t} p$, nos da un método de solución. Nótese también que hemos probado el siguiente teorema de unicidad.

3.5 Teorema. *Si p y q son continuas sobre un intervalo $\mathscr{I}$, si $t_0 \in \mathscr{I}$, y si x_0 es un número real cualquiera, entonces la ecuación diferencial*

$$\dot{x} + px = q$$

tiene una solución única sobre $\mathscr{I}$ que satisface $x(t_0) = x_0$.

3.6 Ejemplo. Resuélvanse

a) $\dot{x} + 3x = \cos$ *b)* $\dot{x}(t) - 2tx(t) = e^{2t}$, $x(0) = 1$.

SOLUCIÓN.

a) Un factor de integración e^{3t}:

$$e^{3t}(\dot{x}(t) + 3x(t)) = D_t(e^{3t}x(t)) = e^{3t}\cos t = D_t\left[\frac{e^{3t}}{10}(3\cos t + \operatorname{sen} t)\right].$$

Por tanto,

$$e^{3t}x(t) = \frac{e^{3t}}{10}(3\cos t + \operatorname{sen} t) + c,$$

y

$$x(t) = ce^{-3t} + \tfrac{3}{10}\cos t + \tfrac{1}{10}\operatorname{sen} t$$

es la solución general.

b) Aquí $p(t) = -2t$ y un factor de integración es e^{-t^2}.

$$e^{-t^2}(\dot{x}(t) - 2tx(t)) = D_t(e^{-t^2}x(t)) = e^{-t^2+2t},$$

$$e^{-t^2}x(t) - x(0) = \int_0^t e^{-s^2+2s}\,ds,$$

y

$$x(t) = e^{t^2} + e^{t^2}\int_0^t e^{-s^2+2^2}\,ds.$$

La integral $\int_0^t e^{-s^2+2s}\,ds$ no puede expresarse en términos de las funciones elementales, pero puede expresarse en términos de una función muy conocida

para la que existen tablas extensas. La función de error, denotada por "fer" está definida por

$$\text{fer } x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du .$$

Ahora bien

$$\begin{aligned}\int_0^t e^{-s^2+2s} ds &= e \int_0^t e^{-(s-1)^2} ds \\ &= e \int_{-1}^{t-1} e^{-u^2} du \\ &= e \int_0^{t-1} e^{-u^2} du + e \int_0^1 e^{-u^2} du \\ &= \frac{e\sqrt{\pi}}{2} [\text{fer } (t-1) + \text{fer } 1].\end{aligned}$$

Por tanto, la solución expresada en términos de la función de error es

$$x(t) = e^{t^2} + \frac{\sqrt{\pi}}{2} e^{t^2+1} [\text{fer } (t-1) + \text{fer } 1].$$

Usando tablas de la función de error, puede calcularse fácilmente una tabla para la solución. En cualquier caso, debe recordarse que hay muchos métodos numéricos para calcular valores de una integral definida y que con las modernas máquinas calculadoras esto es un trabajo de rutina.

3.7 Ejemplo. La razón de decaimiento radiactivo de un elemento se encuentra que es proporcional al número de átomos presentes. Así pues, si $N(t)$ es el número de átomos en el instante t, entonces

$$\dot{N} = -\lambda N;$$

a λ se le llama la constante de decaimiento.[1] El tiempo T requerido para que se desintegre la mitad del número original se llama la vida media del elemento. Pruébese que la vida media T y la constante de decaimiento λ están relacionadas por la fórmula

$$T\lambda = \ln 2.$$

[1] La función N está valuada en los enteros y a menos que sea constante no es continua y ciertamente no tiene derivada. Es, sin embargo, cierto que para un gran número de partículas el proceso puede considerarse continuo más bien que discreto, y la función continua N es un útil tipo de aproximación. Véase R.P. Agnew, *Differential Equations*, II Edic., McGraw-Hill, págs. 85-90.

SOLUCIÓN

$$\dot{N}+\lambda N = 0$$

$$e^{\lambda\tau}(N(\tau)+\lambda N(\tau)) = D_\tau[e^{\lambda\tau}N(\tau)] = 0.$$

Haciendo $N(0)=N_0$, obtenemos, por integración,

$$\int_0^t D_\tau[e^{\lambda\tau}N(\tau)]\,d\tau = e^{\lambda t}N(t)-N_0 = 0.$$

De donde

$$N(t) = N_0 e^{-\lambda t}.$$

Por la definición de T

$$N(T) = \tfrac{1}{2}N_0 = N_0 e^{-\lambda T},$$

y

$$e^{\lambda T} = 2.$$

Por tanto,

$$\lambda T = \ln 2.$$

Problemas

Resuélvanse. (Proporciónense las soluciones en el intervalo o intervalos mayores que se pueda. No hay por qué suponer que siempre será posible expresar las soluciones en términos de funciones elementales.)

1. $\dot{x}+3x = 0$ **2.** $\dot{x}-3x = 0$ **3.** $\dot{x}+bx = 0$

4. $\dot{x}+3x = e^t$ **5.** $\dot{x}+3x = e^{3t}$ **6.** $\dot{x}+3x = e^{-3t}$

7. $\dot{x}+bx = e^{at}$ **8.** $4\dot{x}+x = 1$ **9.** $y' = \dfrac{y-x}{x}$

10. $tx'+x = t^2,\quad x(1) = 3$ **11.** $tx'+x = t^2,\quad x(0) = 0$

12. $tx'+x = t^2,\quad x(0) = 1$ **13.** $x'+tx = t,\quad x(1) = 0$

14. $x'-tx = 0,\quad x(0) = 0$ **15.** $x'-tx = 1,\quad x(0) = 0$

16. $x'-tx = 1,\quad x(0) = 1$ **17.** $x'+tx = 1,\quad x(0) = 2$

18. $\ddot{x}+b\dot{x} = t,\quad x(0) = 1,\ \dot{x}(0) = 0$

19. $\dot{x} = x$
$\dot{y} = x+y,\quad x(0) = c_1$ y $y(0) = c_2$

20. $\dot{x} = \lambda_1 x+y$
$\dot{y} = \lambda_2 y,\quad \lambda_1 \neq \lambda_2,\ x(0) = c_1,\ y(0) = c_2$

21. $\dot{x} = \lambda_1 x+y$
$\dot{y} = \lambda_2 y,\quad \lambda_1 = \lambda_2,\ x(0) = c_1,\ y(0) = c_2$

22. $\dot{x} = \lambda_1 x$
$\dot{y} = x+\lambda_1 y,\quad x(0) = c_1,\ y(0) = c_2$

23. El torio C tiene una vida media de 61 minutos. ¿Qué tiempo transcurrirá para que se desintegre el 90% de torio C?

24. El radiocarbono (C^{14}) se forma en la atmósfera superior por el bombardeo de los rayos cósmicos, entra en los sistemas vivos por un proceso de intercambio, alcanzándose al cabo del tiempo una concentración de equilibrio. El radiocarbono tiene una vida media de 5.6×10^3 años.

a) Una viga de ciprés de la tumba de Sneferu en Egipto contiene el 55% de la cantidad de radiocarbono que tiene la materia viva. Estímese su edad.

b) El carbón vegetal procedente de un árbol que murió a causa de la erupción del volcán que formó el Lago del Cráter en Oregón, contiene el 45% de la cantidad de radiocarbono que tiene la materia viva. Estímese la fecha de la erupción.

25. El torio A se desintegra (dejando libre una partícula alfa) y forma torio B. El torio B se desintegra (liberando una partícula beta) y forma torio C. El torio A tiene una vida media de 0.14 segundos. El torio B tiene una vida media de 10.6 horas. El torio C tiene una vida media de 61 minutos. Si comenzamos con el torio A, ¿qué porcentajes de torio B y C se tendrían después de una hora?

26. En 1701 Newton propuso una ley aproximada para la razón con que un cuerpo cede calor al ambiente que le rodea. Si la capacidad calorífica del cuerpo es una constante, la ley (llamada *ley del enfriamiento de Newton*) nos dice que la temperatura de un cuerpo cambia a una velocidad que es proporcional a la diferencia de temperatura entre el cuerpo y el medio que le rodea. Un termómetro que marca 35.5°C se coloca en la boca de un paciente. Un minuto más tarde la lectura es de 36.6°C, y después de un minuto más se lee 37.5°C. Estímese la temperatura del paciente.

27. Un químico desea enfriar un recipiente hasta 80°. Coloca el recipiente en una gran pila de agua corriente a 45° de temperatura. Al comienzo, la temperatura del contenido del vaso era de 120°. Lo agita constantemente y nota que después de 10 minutos la temperatura ha descendido hasta 100°. Estímese el tiempo total que se necesita para enfriarlo hasta la temperatura deseada.

28. Considérese la ecuación diferencial

$$\dot{x} + cx = f$$

donde c es una constante y f es la función definida por

$$f(t) = \begin{cases} 0, & t \leqslant 0 \\ 1, & 0 < t \leqslant 1 \\ 0, & t > 1 \end{cases}$$

En muchas aplicaciones físicas se sabe que, aparte de satisfacer la ecuación diferencial sobre cada intervalo donde f es continua, x es una función continua. Bajo estas condiciones determínese x dado que $x(0) = 0$. Dibújese la gráfica de la solución (suponiendo c positivo).

29. Supongamos que y es una función positiva y es una solución sobre un intervalo $\mathscr{I}$ de la ecuación diferencial (llamada *ecuación de Bernoulli*)

$$y' + py = qy^n$$

donde n es una constante diferente de 0 y 1. Pruébese que $z = y^{n-1}$ es una solución sobre $\mathscr{I}$ de la ecuación diferencial lineal

$$z' + (1-n)pz = (1-n)q.$$

Resuélvase la ecuación de Bernoulli

$$y'(x) + xy(x) = y^{1/2}(x)\,.$$

30. Si $N(t)$ es el número de individuos de una población en el instante t se sabe que para poblaciones homogéneas aisladas la velocidad de crecimiento de la población es aproximadamente igual a

$$N' = kN(c-N) \quad \text{(ecuación logística de Verhulst-Pearl).}$$

El número c representa la población máxima que el medio puede soportar. Se sabe, de acuerdo con los teoremas de existencia y unicidad, que esta ecuación tiene una solución única que satisface $N(0) = N_0$. Usando el resultado del problema 29, demuéstrese que la solución que satisface la ecuación logística es

$$N(t) = \frac{c}{1 + be^{-kct}}, \quad \text{donde } b = \frac{c}{N(0)} - 1\,.$$

31. Sobre la base de los siguientes conjuntos de datos sobre el crecimiento de la población en los Estados Unidos, predígase la población en 1960.

	Censo	*Población en millones*
a)	1800	5
	1850	25
	1900	76
b)	1860	31
	1890	63
	1920	106
c)	1920	106
	1930	123
	1940	132

32. Un tanque de mezcla de 378.5 litros de capacidad se llena de una salmuera que contiene 18 kilogramos de sal. La solución en el tanque va saliendo a razón de 37.87 litros por segundo y, al mismo tiempo, el tanque se está llenando con una salmuera que contiene 90 gramos de sal por litro. ¿Cuál es la concentración de sal de la salmuera del tanque en el instante t?

4. EXTENSIÓN DE LA FUNCIÓN EXPONENCIAL

A todos nos es familiar la función exponencial como una función real de variable real (es decir, como una función de R en R). En esta sección queremos extender el dominio de definición de la función exponencial del sistema de los números reales R al sistema de los números complejos C. Estudiaremos a continuación algunas de las propiedades fundamentales de esta función exponencial extendida.

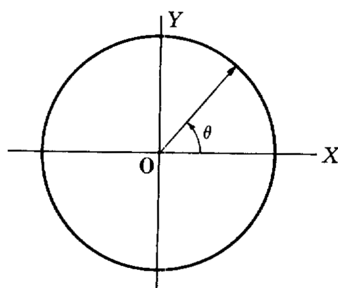
Comenzaremos con una definición.

4.1 Definición

$$e^{i\theta} = \cos\theta + i \operatorname{sen}\theta, \quad \theta \in \mathrm{R}.$$

En el plano complejo $e^{i\theta}$ es el punto $(\cos\theta, \operatorname{sen}\theta)$, o si lo interpretamos como un vector $e^{i\theta}$ es el vector unitario cuyo ángulo de inclinación con el eje real es θ (figura 1). Para decirlo de otra forma, $e^{i\theta}$ es el número complejo cuyo valor absoluto es 1 y cuyo argumento es θ.

Es entonces una simple consecuencia de las fórmulas de adición para las funciones seno y coseno[1] que



$e^{i\theta} = (\cos\theta, \operatorname{sen}\theta)$

FIGURA 1

[1] En el capítulo 13, sección 4, la función exponencial e^z, con $z = x + iy$ es un número complejo, se define independientemente de las funciones trigonométricas. Se establece

4.2
$$\begin{aligned} e^{i\theta_1} e^{i\theta_2} &= (\cos\theta_1 + i \operatorname{sen}\theta_1)(\cos\theta_2 + i \operatorname{sen}\theta_2) \\ &= (\cos\theta_1 \cos\theta_2 - \operatorname{sen}\theta_1 \operatorname{sen}\theta_2) \\ &\quad + i(\cos\theta_1 \operatorname{sen}\theta_2 + \cos\theta_2 \operatorname{sen}\theta_1) \\ &= \cos(\theta_1+\theta_2) + i \operatorname{sen}(\theta_1+\theta_2) = e^{i(\theta_1+\theta_2)}. \end{aligned}$$

Vemos también que esta ley de exponentes implica las fórmulas de adición de la trigonometría y la simplicidad de operar con la exponencial compleja hace ventajoso trabajar con la función exponencial compleja en lugar de con las funciones trigonométricas.

Antes de ilustrar esto con varios ejemplos, nótese que es una consecuencia directa de la definición de $e^{i\theta}$ que

4.3
$$\cos\theta = \operatorname{Rl}(e^{i\theta}) = \frac{e^{i\theta}+e^{-i\theta}}{2}$$

y

4.4
$$\operatorname{sen}\theta = \operatorname{Im}(e^{i\theta}) = \frac{e^{i\theta}-e^{-i\theta}}{2i}.$$

4.5 Ejemplo. Pruébese que $2 \operatorname{sen} a \cos b = \operatorname{sen}(a+b) + \operatorname{sen}(a-b)$.

SOLUCIÓN.

$$\begin{aligned} 2 \operatorname{sen} a \cos b &= 2\, \frac{e^{ia}-e^{-ia}}{2i}\, \frac{e^{ib}+e^{-ib}}{2} \\ &= \frac{1}{2i}\left[e^{i(a+b)} - e^{-i(a+b)} + e^{i(a-b)} - e^{-i(a-b)}\right] \\ &= \operatorname{sen}(a+b) + \operatorname{sen}(a-b). \end{aligned}$$

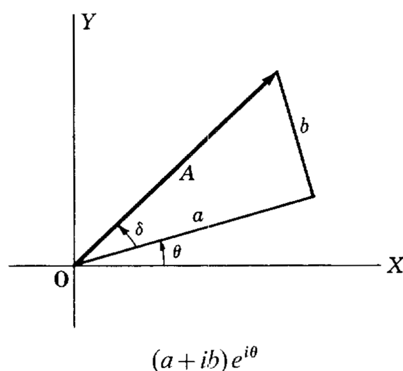
---

la ley de exponentes $e^{z_1}e^{z_2} = e^{z_1+z_2}$, y se definen las funciones trigonométricas seno y coseno mediante

$$\cos z = \frac{e^{iz}+e^{-iz}}{2}$$

$$\operatorname{sen} z = \frac{e^{iz}-e^{-iz}}{2i}.$$

Las fórmulas de adición para el seno y el coseno resultan una consecuencia de la ley de los exponentes.



$(a+ib)e^{i\theta}$

FIGURA 2

4.6 Ejemplo. Pruébese que:

$$a \operatorname{sen} \theta + b \cos \theta = A \operatorname{sen} (\theta + \delta),$$

donde $a+ib = Ae^{i\delta}$; es decir, $A = \sqrt{a^2+b^2}$ y $\delta = \operatorname{Arg}(a+ib)$.

SOLUCIÓN. (Figura 2.)

$$\begin{aligned}(a+ib)e^{i\theta} &= (a+ib)(\cos \theta + i \operatorname{sen} \theta)\\ &= (a \cos \theta - b \operatorname{sen} \theta) + i(a \operatorname{sen} \theta + b \cos \theta).\end{aligned}$$

Además,

$$(a+ib)e^{i\theta} = Ae^{i\delta}e^{i\theta} = Ae^{i(\theta+\delta)} = A \cos(\theta+\delta) + iA \operatorname{sen}(\theta+\delta).$$

Por tanto

$$A \cos(\theta+\delta) + iA \operatorname{sen}(\theta+\delta) = (a \cos \theta - b \operatorname{sen} \theta) + i(a \operatorname{sen} \theta + b \cos \theta).$$

Igualando las partes reales e imaginarias, tenemos

$$\begin{aligned}a \cos \theta - b \operatorname{sen} \theta &= A \cos(\theta+\delta)\\ a \operatorname{sen} \theta + b \cos \theta &= A \operatorname{sen}(\theta+\delta).\end{aligned}$$

Definimos ahora e^z para cualquier número complejo $z = x+iy$.

4.7 Definición. $e^z = e^x e^{iy} = e^x(\cos y + i \operatorname{sen} y)$ *para todo* $z = x+iy$ *en* C.

La función exponencial compleja exp definida por $\exp z = e^z$ es ahora una función de C en C. Cuando restringimos z a R (es decir, cuando $y = 0$), esta función exponencial se reduce a la función exponencial real de variable real.

Las siguientes propiedades se derivan fácilmente:

4.8 $\qquad e^{z_1}e^{z_2} = e^{z_1+z_2}$ para cualesquier $z_1, z_2 \in C$.

4.9 $e^z \neq 0$; y $(e^z)^{-1} = e^{-z}$ para todo $z \in C$.

4.10 $e^z = 1 \Leftrightarrow z = 2\pi n i$ para algún entero n.

4.11 $e^{z_1} = e^{z_2} \Leftrightarrow z_1 = z_2 + 2\pi n i$ para algún entero n.

En nuestro estudio de las ecuaciones diferenciales estamos principalmente interesados en la función f definida por $f(t) = e^{\lambda t}$, donde λ es un número complejo y t es un número real. La función f es una función compleja de variable real; una función de R en C. Denotemos por f cualquier función de R en C, y sea $f = u + iv = (u, v)$ $(f(t) = u(t) + iv(t))$. Entonces, f puede considerarse también como una función de R en R^2, y el límite y la derivada de f —que fueron definidas para funciones de R en R^2— están definidas para f. Así pues, $f = u + iv$ es continua sobre $\mathscr{I}$ si y sólo si u y v son continuas sobre $\mathscr{I}$. La derivada de f es $f' = u' + iv'$.

4.12 Definición. *Si $f = u + iv$ es una función de* R *en C y si u y v son integrables sobre $\mathscr{I}$, definimos entonces para cada a, $b \in \mathscr{I}$:*

$$\int_a^b f = \int_a^b u + i \int_a^b v.$$

Es entonces evidente que *el teorema fundamental del cálculo se verifica para funciones de* R *en C* (para funciones complejas de una variable real).

Definamos f por $f(t) = e^{\lambda t}$ para toda $t \in$ R y un cierto número complejo $\lambda = \alpha + i\beta$. Entonces

$$f(t) = e^{\lambda t} = e^{\alpha t}(\cos \beta t + i \text{ sen } \beta t)$$

y

$$\begin{aligned} f'(t) &= D_t(e^{\lambda t}) = \alpha e^{\alpha t}(\cos \beta t + i \text{ sen } \beta t) + \beta e^{\alpha t}(-\text{ sen } \beta t + i \cos \beta t) \\ &= \alpha e^{\alpha t} e^{i\beta t} + i\beta e^{\alpha t} e^{i\beta t} \\ &= (\alpha + i\beta) e^{(\alpha + i\beta)t} = \lambda e^{\lambda t}. \end{aligned}$$

De donde

4.13 $$D_t(e^{\lambda t}) = \lambda e^{\lambda t}.$$

Como el teorema fundamental del cálculo se verifica para funciones de R en C, el teorema 2.1, pág. 591, se verifica si y_0 es un número complejo cualquiera y si f es una función continua de R en C. Además, como $D_t(e^{\lambda t}) = \lambda e^{\lambda t}$ donde λ es un número complejo, los resultados obtenidos en la sección 3 para la ecuación lineal de primer orden $x' + px = q$ se verifican cuando p y q son funciones complejas.

4.14 Ejemplo. Determínese $\int e^{at} \cos bt\, dt$, a y b reales con $a^2 + b^2 \neq 0$.

SOLUCIÓN. De acuerdo con 4.13

$$\int e^{(a+ib)t}\,dt = \frac{1}{a+ib}\,e^{(a+ib)t}.$$

Igualando las partes reales y las partes imaginarias, obtenemos

$$\int e^{at}\cos bt\,dt = \mathrm{Rl}\,\frac{1}{a+ib}\,e^{(a+ib)t}$$

$$= \frac{e^{at}}{a^2+b^2}(a\cos bt + b\,\mathrm{sen}\,bt)$$

e

$$\int e^{at}\,\mathrm{sen}\;bt\,dt = \mathrm{Im}\,\frac{1}{a+ib}\,e^{(a+ib)t}$$

$$= \frac{e^{at}}{a^2+b^2}(a\,\mathrm{sen}\,bt - b\cos bt).$$

4.15 Ejemplo. Resuélvase

$$\dot{x} = ax - by$$
$$\dot{y} = bx + ay.$$

SOLUCIÓN. Sea $z = x + iy$. Entonces x y y son soluciones si y sólo si

$$\begin{aligned}\dot{z} &= \dot{x} + i\dot{y} = (ax-by) + i(bx+ay)\\ &= (a+ib)x + (a+ib)iy\\ &= (a+ib)z.\end{aligned}$$

De donde

$$z(t) = Ce^{(a+ib)t}.$$

Haciendo $C = C_1 + iC_2$, obtenemos

$$x(t) = e^{at}(C_1\cos bt - C_2\,\mathrm{sen}\,bt)$$

$$y(t) = e^{at}(C_1\,\mathrm{sen}\,bt + C_2\cos bt).$$

O, haciendo $C = Ae^{i\delta}$, obtenemos

$$x(t) = \mathrm{Rl}\,[Ae^{i\delta}e^{(a+ib)t}] = Ae^{at}\cos(bt+\delta)$$

$$y(t) = \mathrm{Im}\,[Ae^{i\delta}e^{(a+ib)t}] = Ae^{at}\,\mathrm{sen}\,(bt+\delta).$$

Problemas

1. Pruébese que

a) $2 \text{ sen } a \text{ sen } b = \cos(a-b) - \cos(a+b)$

b) $\text{sen}^3 \theta = \frac{1}{4}(3 \text{ sen } \theta - \text{sen } 3\theta)$

c) $\cos^4 \theta = \frac{1}{8}(3+4\cos 2\theta + \cos 4\theta)$.

2. Exprésense

a) $2\cos\theta - 2\text{ sen }\theta$

b) $5\text{ sen }\theta + 4\cos\theta$

en la forma $A \text{ sen } (\theta+\delta)$ Obténganse A y δ tanto analítica como gráficamente.

3. Pruébese que:

a) 4.8 b) 4.9 c) 4.10 d) 4.11.

4. Demuéstrese que

a) $$\int_0^{2\pi} e^{im\theta} e^{-in\theta}\, d\theta = \begin{cases} 0 & \text{si } m \neq n \\ 2\pi & \text{si } m = n, \ m \text{ y } n \text{ enteros} \end{cases}$$

b) $$\int_0^{2\pi} \text{sen } m\theta \text{ sen } n\theta\, d\theta = \int_0^{2\pi} \cos m\theta \cos n\theta\, d\theta = \begin{cases} 0 & \text{si } m \neq n \\ \pi & \text{si } m = n \neq 0 \end{cases}$$

c) $$\int_0^{2\pi} \cos m\theta \text{ sen } n\theta\, d\theta = 0, \quad m \text{ y } n \text{ enteros.}$$

5. Si λ es un número complejo cualquiera, pruébese que $x(t) = e^{\lambda t}$ es la única solución de

$$\dot{x} = \lambda x \qquad x(0) = 1.$$

6. Pruébese que $x(t) = e^{\lambda t}$ es una solución de

$$D^n x + a_1 D^{n-1} x + \ldots + a_n x = 0$$

si y sólo si λ es una raíz de

$$\varphi(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \ldots + a_n = 0 .$$

7. Pruébese que $x(t) = c_1 e^{\lambda t} + c_2 e^{\lambda_2 t}$ es una solución de

$$\ddot{x} - (\lambda_1 + \lambda_2)\dot{x} + \lambda_1 \lambda_2 x = 0 .$$

8. Pruébese que

a) Si $\lambda_1 \neq \lambda_2$, entonces $c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} = 0$ para todo $t \in \langle -\infty, \infty \rangle$ implica $c_1 = c_2 = 0$.

b) $c_1 e^{\lambda_1 t} + c_2 t e^{\lambda_1 t} = 0$ para todo $t \in \langle -\infty, \infty \rangle$ implica $c_1 = c_2 = 0$.

(Funciones con esta propiedad se dice que son *linealmente independientes* sobre, en este caso, el intervalo $\langle -\infty, \infty \rangle$.)

9. Si b, c, ω y A son números reales y $z = x + iy$ es una solución de

$$z'' + bz' + cz = Ae^{i\omega t}$$

pruébese que $x = \text{Rl}\,(z)$ es una solución de

$$x'' + bx' + cx = A \cos \omega t$$

y $y = \text{Im}\,(z)$ es una solución de

$$y'' + by' + cy = A \operatorname{sen} \omega t\ .$$

10. En cada una de las esquinas de una mesita se posa una mosca. Las moscas miran hacia el interior y comienzan a moverse al mismo tiempo y caminan a la misma velocidad, cada una de ellas hacia la posición que en cada momento ocupa la que está a su izquierda. ¿Qué camino recorre cada mosca ?, ¿cuánto caminan antes de encontrarse ?

11. Derívense las fórmulas ($\theta \neq 2m\pi$)

a) $$\sum_{k=0}^{n} \cos k\theta = 1 + \cos\theta + \cdots + \cos n\theta = \frac{1}{2} + \frac{\operatorname{sen}(n+\frac{1}{2})\theta}{2\operatorname{sen}\frac{1}{2}\theta}$$

b) $$\sum_{k=0}^{n} \operatorname{sen} k\theta = \frac{\cos\frac{1}{2}\theta - \cos(n+\frac{1}{2})\theta}{2\operatorname{sen}\frac{1}{2}\theta}\ .$$

5. SISTEMAS LINEALES BIDIMENSIONALES. COEFICIENTES CONSTANTES

Queremos estudiar aquí un par de ecuaciones diferenciales lineales de primer orden

5.1 $$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 \end{aligned}$$

en las dos incógnitas x_1 y x_2. Los coeficientes a_{11}, a_{12}, a_{21}, a_{22} son números reales o complejos dados (constantes). Hay dos casos en que la solución general puede formularse inmediatamente. Supongamos que las ecuaciones tienen la sencilla forma

$$\begin{aligned} \dot{x}_1 &= \lambda_1 x_1 \\ \dot{x}_2 &= \lambda_2 x_2\ . \end{aligned}$$

Entonces, las dos ecuaciones son independientes una de otra y la solución

general es $x_1(t) = c_1 e^{\lambda_1 t}$, $x_2(t) = c_2 e^{\lambda_2 t}$. El otro caso que es sencillo resolver, es cuando $a_{12} = 0$ y las ecuaciones son de la forma

$$\begin{aligned}\dot{x}_1 &= \lambda_1 x_1 \\ \dot{x}_2 &= \lambda_2 x_2 + a_{21} x_1 .\end{aligned}$$

El problema entonces es resolver sucesivamente un par de ecuaciones de primer orden:

$$\begin{aligned}x_1(t) &= c_1 e^{\lambda_1 t} \\ \dot{x}_2(t) &= \lambda_2 x_2(t) + a_{21} c_1 e^{\lambda_1 t}\end{aligned}$$

No formularemos en este momento la solución de x_2, sino que nos limitaremos a señalar que ésta es una ecuación de las que sabemos cómo resolver.

La forma en que resolveremos el sistema general 5.1 será mediante un cambio de coordenadas que lo reduzca a una de las formas sencillas que acabamos de mencionar. Nuestra discusión se simplifica en gran parte si introducimos la notación matricial. En la sección 3 del capítulo 5 consideramos matrices de números reales. Aquí tomaremos como sistema de números el campo de los números complejos C y nos ocuparemos, por tanto, de matrices de números complejos. Es fácil ver que estas matrices tienen las propiedades que probamos en el capítulo 5 para matrices de números reales.

Sean

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{y} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix};$$

A es una matriz 2×2 de números complejos y $\mathbf{x}$ es una función con dominio en R y rango en C^2, el espacio vectorial bidimensional sobre el campo complejo (capítulo 1, sección 8). Podemos escribir entonces 5.1 en la forma

5.4 $$\dot{\mathbf{x}} = A\mathbf{x} \qquad \dot{\mathbf{x}}((t) = A\mathbf{x}(t)).$$

Así pues, en la notación matricial el sistema de ecuaciones diferenciales 5.1 se transforma en la ecuación diferencial vectorial de primer orden simple 5.4, donde $\mathbf{y} = A\mathbf{x}$ es la función vectorial cuyo valor en t es el vector $\mathbf{y}(t) = A\mathbf{x}(t)$ en C^2.

La matriz A define una función cuyo dominio es C^2 y cuyo rango está en C^2: para cualquier $\mathbf{v} \in C^2$, $A\mathbf{v}$ corresponde a $\mathbf{v}$. Tal función se llama *transformación de* C^2. Esta transformación tiene la propiedad de que

$$A(c_1 \mathbf{v}^1 + c_2 \mathbf{v}^2) = c_1 A\mathbf{v}^1 + c_2 A\mathbf{v}^2$$

para todos los números complejos c_1, c_2 y todos los vectores $\mathbf{v}^1$, $\mathbf{v}^2$ en C^2. Una transformación con esta propiedad se dice que es *lineal.* Es fácil probar que la linealidad de la transformación asociada con A implica que *cualquier combinación lineal de soluciones de* 5.4 *es una solución:* sean $\mathbf{x}^1$

y $\mathbf{x}^2$ soluciones de 5.4 y sean c_1 y c_2 un par de constantes cualesquiera. Entonces $\mathbf{x} = c_1\mathbf{x}^1 + c_2\mathbf{x}^2$ es una solución de 5.4 ya que

$$\begin{aligned}\dot{\mathbf{x}}(t) &= c_1\dot{\mathbf{x}}^1(t) + c_2\dot{\mathbf{x}}^2(t) = c_1 A\mathbf{x}^1(t) + c_2 A\mathbf{x}^2(t)\\ &= A(c_1\mathbf{x}^1(t) + c_2\mathbf{x}^2(t)) = A\mathbf{x}(t).\end{aligned}$$

Nuestro objetivo es encontrar soluciones de la ecuación 5.4, que es el sistema 5.1 en la notación matricial. Resulta, como veremos dentro de poco, que el problema de resolver el sistema lineal de ecuaciones diferenciales puede resolverse como si fuera un sistema de álgebra lineal. El problema algebraico es el de resolver

5.5 $$A\mathbf{v} = \lambda\mathbf{v}.$$

Queremos encontrar un vector no trivial (no cero) $\mathbf{v}$ y un número λ que satisfagan 5.5. El número λ se llama *valor característico* de la matriz A, y el vector $\mathbf{v}$ un *vector característico* de A ($\mathbf{v} \neq \mathbf{0}$). El número λ se llama también *valor propio* o *eigenvalor* de A, y $\mathbf{v}$ se llama *vector propio* o *eigenvector* de A. Si escribimos esta ecuación $A\mathbf{v} = \lambda\mathbf{v}$ en términos de sus componentes, tenemos

$$\begin{aligned}a_{11}v_1 + a_{12}v_2 &= \lambda v_1\\ a_{21}v_1 + a_{22}v_2 &= \lambda v_2.\end{aligned}$$

Queremos, por tanto, encontrar soluciones no triviales de las ecuaciones lineales homogéneas

5.6 $$\begin{aligned}(a_{11}-\lambda)v_1 + a_{12}v_2 &= 0\\ a_{21}v_1 + (a_{22}-\lambda)v_2 &= 0.\end{aligned}$$

Sabemos que hay soluciones no triviales si y sólo si

$$\begin{vmatrix} a_{11}-\lambda & a_{12}\\ a_{21} & a_{22}-\lambda\end{vmatrix} = (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21} = 0.$$

Así pues, los valores característicos son los ceros del polinomio

$$\varphi(\lambda) = \lambda^2 - (a_{11}+a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}).$$

A este polinomio se le llama *polinomio característico* de A. $\varphi(\lambda) = 0$ se llama *ecuación característica* de A, y los valores característicos λ_1 y λ_2 se llaman también *raíces características* de A. Sean λ_1 y λ_2 las raíces características de A. Entonces $\varphi(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)$ y

$$\lambda_1 + \lambda_2 = a_{11} + a_{22}, \quad \lambda_1\lambda_2 = a_{11}a_{22} - a_{12}a_{21}.$$

Si $a_{12} = a_{21} = 0$, entonces por 5.6 vemos que $\lambda_1 = a_{11}$, $\lambda_2 = a_{22}$, y los vectores característicos correspondientes son $\mathbf{v}^1 = \begin{pmatrix}1\\0\end{pmatrix}$, $\mathbf{v}^2 = \begin{pmatrix}0\\1\end{pmatrix}$. En este caso, como ya hemos señalado, podemos hallar la solución general de la ecuación diferencial 5.1 inmediatamente. Por tanto, excluyendo este caso

trivial, podemos suponer —cambiando los subíndices si es necesario— que $a_{12} \neq 0$. Entonces, correspondiéndose con el valor característico λ_1 tenemos el vector característico $\mathbf{v}^1 = \begin{pmatrix} a_{12} \\ \lambda_1 - a_{11} \end{pmatrix}$, y correspondiéndose con el valor característico λ_2 tenemos el vector característico $\mathbf{v}^2 = \begin{pmatrix} a_{12} \\ \lambda_2 - a_{11} \end{pmatrix}$. Si $\lambda_1 \neq \lambda_2$, es fácil ver que $\mathbf{v}^1$ y $\mathbf{v}^2$ son linealmente independientes.

Volviendo a nuestra ecuación diferencial 5.1 o 5.4, suponemos, en primer término, que los valores característicos de A son distintos ($\lambda_1 \neq \lambda_2$) y que $a_{12} \neq 0$. Entonces $\mathbf{v}^1$ y $\mathbf{v}^2$ son linealmente independientes y todo vector $\mathbf{x}(t)$ es expresable de modo único en la forma

$$\mathbf{x}(t) = y_1(t)\mathbf{v}^1 + y_2(t)\mathbf{v}^2.$$

Esto es un cambio de coordenadas con los nuevos ejes de coordenadas en las direcciones $\mathbf{v}^1$ y $\mathbf{v}^2$. Si $\mathbf{x} = y_1\mathbf{v}^1 + y_2\mathbf{v}^2$ es una solución de 5.1, entonces

$$\begin{aligned}\dot{\mathbf{x}} &= \dot{y}_1\mathbf{v}^1 + \dot{y}_2\mathbf{v}^2 = A(y_1\mathbf{v}^1 + y_2\mathbf{v}^2) = y_1 A\mathbf{v}^1 + y_2 A\mathbf{v}^2 \\ &= \lambda_1 y_1\mathbf{v}^1 + \lambda_2 y_2\mathbf{v}^2.\end{aligned}$$

Por tanto, como $\mathbf{v}^1$ y $\mathbf{v}^2$ son linealmente independientes

$$\begin{aligned}\dot{y}_1 &= \lambda_1 y_1 \\ \dot{y}_2 &= \lambda_2 y_2.\end{aligned}$$

El cambio de coordenadas nos lleva a este sencillo sistema cuya solución es

$$y_1(t) = c_1 e^{\lambda_1 t}, \quad y_2(t) = c_2 e^{\lambda_2 t}.$$

Hemos, por tanto, probado que si $\lambda_1 \neq \lambda_2$ y $a_{12} \neq 0$, toda solución de 5.1 es de la forma

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{v}^1 + c_2 e^{\lambda_2 t}\mathbf{v}^2$$

donde $\mathbf{v}^i = \begin{pmatrix} a_{12} \\ \lambda_i - a_{11} \end{pmatrix}$. La solución general puede también escribirse

5.7
$$\begin{aligned}x_1(t) &= a_{12}c_1 e^{\lambda_1 t} + a_{12}c_2 e^{\lambda_2 t} \\ x_2(t) &= (\lambda_1 - a_{11})c_1 e^{\lambda_1 t} + (\lambda_2 - a_{11})c_2 e^{\lambda_2 t}.\end{aligned}$$

Recíprocamente se verifica con toda facilidad que cada función de esta forma es una solución. Por tanto 5.7 es la solución general de 5.1 cuando $\lambda_1 \neq \lambda_2$ y $a_{12} \neq 0$.

5.8 Ejemplo. Resuélvase

$$\begin{aligned}\dot{x} &= x + y \\ \dot{y} &= x - y.\end{aligned}$$

SOLUCIÓN. Aquí $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, y su ecuación característica es

$$\begin{vmatrix} 1-\lambda & 1 \\ 1 & -1-\lambda \end{vmatrix} = \lambda^2 - 2 = 0.$$

Por tanto, $\lambda_1 = \sqrt{2}$, $\mathbf{v}^1 = \begin{pmatrix} 1 \\ \sqrt{2}-1 \end{pmatrix}$, $\lambda_2 = -\sqrt{2}$, $\mathbf{v}^2 = \begin{pmatrix} 1 \\ -\sqrt{2}-1 \end{pmatrix}$, y la solución general es

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 e^{\sqrt{2}t} \begin{pmatrix} 1 \\ \sqrt{2}-1 \end{pmatrix} + c_2 e^{-\sqrt{2}t} \begin{pmatrix} 1 \\ -\sqrt{2}-1 \end{pmatrix}.$$

De donde

$$\begin{aligned} x(t) &= c_1 e^{\sqrt{2}t} + c_2 e^{-\sqrt{2}t} \\ y(t) &= (\sqrt{2}-1) c_1 e^{\sqrt{2}t} - (\sqrt{2}+1) c_2 e^{-\sqrt{2}t} \end{aligned}$$

es la solución general.

Veamos ahora lo que sucede cuando $\lambda_1 = \lambda_2 = \lambda$ y $a_{12} \neq 0$. Aquí los vectores $\mathbf{v} = \begin{pmatrix} a_{12} \\ \lambda - a_{11} \end{pmatrix}$ y $\mathbf{u} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ son linealmente independientes. Efectuando el cambio de coordenadas

$$\mathbf{x}(t) = y_1(t)\mathbf{v} + y_2(t)\mathbf{u}.$$

Entonces, si x es una solución de 5.1,

$$\begin{aligned} \dot{\mathbf{x}} &= \dot{y}_1 \mathbf{v} + \dot{y}_2 \mathbf{u} = A(y_1 \mathbf{v} + y_2 \mathbf{u}) = y_1 A\mathbf{v} + y_2 A\mathbf{u} \\ &= \lambda y_1 \mathbf{v} + y_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}. \end{aligned}$$

Como $\lambda = \lambda_1 = \lambda_2$, $2\lambda = a_{11} + a_{22}$ y

$$\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \begin{pmatrix} a_{12} \\ 2\lambda - a_{11} \end{pmatrix} = \begin{pmatrix} a_{12} \\ \lambda - a_{11} \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda \end{pmatrix} = \mathbf{v} + \lambda \mathbf{u}.$$

De donde

$$\dot{y}_1 \mathbf{v} + \dot{y}_2 \mathbf{u} = (\lambda y_1 + y_2)\mathbf{v} + \lambda y_2 \mathbf{u},$$

y nuevamente, de acuerdo con la independencia lineal de $\mathbf{v}$ y $\mathbf{u}$ tenemos, en las nuevas coordenadas, el sistema de ecuaciones diferenciales

$$\begin{aligned} \dot{y}_1 &= \lambda y_1 + y_2 \\ \dot{y}_2 &= \lambda y_2. \end{aligned}$$

De donde $y_2(t) = c_1 e^{\lambda t}$, y $\dot{y}_1(t) = \lambda y_1(t) + c_1 e^{\lambda t}$. Sabemos cómo resolver esta ecuación diferencial lineal de primer orden para y_1. Haciéndolo así

obtenemos $y_1(t) = c_1 te^{\lambda t} + c_2 e^{\lambda t}$. De donde, la solución general de 5.1, es

$$\mathbf{x}(t) = (c_1 t + c_2)e^{\lambda t}\mathbf{v} + c_1 e^{\lambda t}\mathbf{u};$$

es decir

5.9
$$x_1(t) = a_{12}(c_1 t + c_2)e^{\lambda t}$$
$$x_2(t) = (\lambda - a_{11})(c_1 t + c_2)e^{\lambda t} + c_1 e^{\lambda t}$$

es la solución general cuando $\lambda_1 = \lambda_2 = \lambda$ y $a_{12} \neq 0$.

5.10 Ejemplo. Resuélvase

$$\dot{x} = x - y$$
$$\dot{y} = x + 3y.$$

SOLUCIÓN. La ecuación característica es

$$\begin{vmatrix} 1-\lambda & -1 \\ 1 & 3-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2 .$$

Este es, entonces, un caso de raíces iguales: $\lambda_1 = \lambda_2 = 2$. De acuerdo con 5.9 la solución general es

$$x(t) = -(c_1 t + c_2)c^{2t}$$
$$y(t) = (c_1 t + c_2 + c_1)c^{2t}.$$

Sumario. *La solución general de*

$$\dot{x}_1 = a_{11}x_1 + a_{12}x_2$$
$$\dot{x}_2 = a_{21}x_1 + a_{22}x_2, \quad a_{12} \neq 0$$

es:

5.7′
$$x_1(t) = a_{12}c_1 e^{\lambda_1 t} + a_{12}c_2 e^{\lambda_2 t}$$
$$x_2(t) = (\lambda_1 - a_{11})c_1 e^{\lambda_1 t} + (\lambda_2 - a_{11})c_2 e^{\lambda_2 t}, \ si\ \lambda_1 \neq \lambda_2 .$$

5.9′
$$x_1(t) = a_{12}(c_1 t + c_2)e^{\lambda t}$$
$$x_2(t) = (\lambda - a_{11})(c_1 t + c_2)e^{\lambda t} + c_1 e^{\lambda t}, \ si\ \lambda_1 = \lambda_2 = \lambda .$$

Cambiando los índices, si es necesario, el único caso restante es $\dot{x}_1 = \lambda_1 x_1$ y $\dot{x}_2 = \lambda_2 x_2$ y en este caso la solución general es $x_1(t) = c_1 e^{\lambda_1 t}$, $x_2(t) = c_2 e^{\lambda_2 t}$. Así pues, hemos encontrado la solución general sobre $\langle -\infty, \infty \rangle$ de 5.1 y vemos que *la solución de los sistemas lineales bidimensionales* $\dot{\mathbf{x}} = A\mathbf{x}$ *con coeficientes constantes puede reducirse al problema algebraico de resolver el problema del valor característico* $A\mathbf{v} = \lambda\mathbf{v}$. Esta afirmación se generaliza a los sistemas n-dimensionales de ecuaciones diferenciales lineales con coeficientes constantes. Las dificultades para el cálculo aumentan tremen-

damente con el aumento de dimensión, pero las ideas matemáticas son las mismas.

5.11 Ejemplo. Resuélvase

$$\begin{aligned} \dot{x} &= x+y+2z \\ \dot{y} &= 2y+2z \\ \dot{z} &= x-y. \end{aligned}$$

SOLUCIÓN. El problema del valor característico es

$$\begin{aligned} x+y+2z &= \lambda x \\ 2y+2z &= \lambda y \\ x-y &= \lambda z \end{aligned}$$

o bien

$$\begin{aligned} (1-\lambda)x+y+2z &= 0 \\ (2-\lambda)y+2z &= 0 \\ x-y-\lambda z &= 0. \end{aligned}$$

Para que existan soluciones no triviales el determinante de los coeficientes debe ser cero:

$$\begin{vmatrix} 1-\lambda & 1 & 2 \\ 0 & 2-\lambda & 2 \\ 1 & -1 & -\lambda \end{vmatrix} = -(\lambda^3-3\lambda^2+2\lambda) = 0.$$

Los valores característicos son 0, 1, 2. Los vectores característicos correspondientes son

$$\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Hagamos el cambio de coordenadas

$$\mathbf{r} = \bar{x}\mathbf{u}+\bar{y}\mathbf{v}+\bar{z}\mathbf{w},$$

donde $\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$; es decir,

$$\begin{aligned} x &= \bar{x}+\bar{y}+\bar{z} \\ y &= \bar{x}+2\bar{y}+\bar{z} \\ z &= -\bar{x}-\bar{y}. \end{aligned}$$

Ahora $\mathbf{r}$ es una solución del sistema si y sólo si

$$\dot{\mathbf{r}} = \dot{\bar{x}}\mathbf{u}+\dot{\bar{y}}\mathbf{v}+\dot{\bar{z}}\mathbf{w} = A\mathbf{r} = \bar{x}A\mathbf{u}+\bar{y}A\mathbf{v}+\bar{z}A\mathbf{w} = 0\bar{x}\mathbf{u}+\bar{y}\mathbf{v}+2\bar{z}\mathbf{w},$$

donde $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 1 & -1 & 0 \end{pmatrix}$. (Los números 0, 1, 2 son los valores característicos de A y $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ son vectores característicos correspondientes : $A\mathbf{u} = 0\mathbf{u}$, $A\mathbf{v} = \mathbf{v}$, $A\mathbf{w} = 2\mathbf{w}$.) De donde

$$\dot{\bar{x}} = 0, \quad \dot{\bar{y}} = \bar{y}, \quad \dot{\bar{z}} = 2\bar{z},$$

y la solución general es

$$\bar{x}(t) = c_1, \quad \bar{y}(t) = c_2 e^t, \quad \bar{z}(t) = c_3 e^{2t}.$$

En términos de las coordenadas originales, la solución general es

$$\begin{aligned} x(t) &= c_1 + c_2 e^t + c_3 e^{2t} \\ y(t) &= c_1 + 2c_2 e^t + c_3 e^{2t} \\ z(t) &= -c_1 - c_2 e^t . \end{aligned}$$

Ahora que conocemos la solución general de 5.1 es fácil probar que:

5.12 Teorema. *Dado un número real t_0 y cualquier vector constante $\mathbf{x}^0$, la ecuación $\dot{\mathbf{x}} = A\mathbf{x}$ tiene una solución única sobre $\langle -\infty, \infty \rangle$ que satisface $\mathbf{x}(t_0) = \mathbf{x}^0$.*

PRUEBA. Vamos a probar el resultado sólo para el caso de valores característicos distintos $\lambda_1 \neq \lambda_2$ y $a_{12} \neq 0$. Como c_1 y c_2 son constantes arbitrarias, podemos escribir la solución general en la forma

$$\mathbf{x}(t) = c_1 e^{\lambda_1(t-t_0)}\mathbf{v}^1 + c_2 e^{\lambda_2(t-t_0)}\mathbf{v}^2 .$$

Queremos encontrar las soluciones que satisfacen $\mathbf{x}(t_0) = \mathbf{x}^0$; es decir, queremos encontrar c_1 y c_2 que satisfagan

$$c_1 \mathbf{v}^1 + c_2 \mathbf{v}_2 = \mathbf{x}^0.$$

Pero $\mathbf{v}^1$ y $\mathbf{v}^2$ son linealmente independientes, y esta ecuación algebraica tiene soluciones únicas. Por tanto, la ecuación diferencial tiene una solución única que satisface $\mathbf{x}(t_0) = \mathbf{x}^0$. Un argumento análogo puede usarse cuando $\lambda_1 = \lambda_2$.

5.13 Ejemplo. Determínese la solución de

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -2x - 3y \end{aligned}$$

que satisface $x(0) = 0$, $y(0) = 1$.

SOLUCIÓN. La ecuación característica es

$$\begin{vmatrix} -\lambda & 1 \\ -2 & -3-\lambda \end{vmatrix} = \lambda^2+3\lambda+2 = 0,$$

y las raíces características son $\lambda_1 = -1$ y $\lambda_2 = -2$. La solución general es

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 e^{-t}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 e^{-2t}\begin{pmatrix} 1 \\ -2 \end{pmatrix},$$

o bien

$$x(t) = c_1 e^{-t}+c_2 e^{-2t}$$

$$y(t) = -c_1 e^{-t}-2c_2 e^{-2t}.$$

Queremos que

$$x(0) = c_1+c_2 = 0$$
$$y(0) = -c_1-2c_2 = 1.$$

De donde $c_1 = 1$, $c_2 = -1$, y la solución requerida es

$$x(t) = e^{-t}-e^{-2t}$$

$$y(t) = -e^{-t}+2e^{-2t}$$

5.14 Ejemplo. Determínese la solución de

$$\dot{x} = y$$
$$\dot{y} = -5x+2y$$

que satisface $x(0) = 1$, $y(0) = 0$.

SOLUCIÓN. La ecuación característica es

$$\begin{vmatrix} -\lambda & 1 \\ -5 & 2-\lambda \end{vmatrix} = \lambda^2-2\lambda+5 = 0;$$

$\lambda_1 = 1+2i$, $\lambda_2 = 1-2i$. La solución general es

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 e^{(1+2i)t}\begin{pmatrix} 1 \\ 1+2i \end{pmatrix} + c_2 e^{(1-2i)t}\begin{pmatrix} 1 \\ 1-2i \end{pmatrix}$$

o bien

$$x(t) = c_1 e^{(1+2i)t}+c_2 e^{(1-2i)t}$$

$$y(t) = (1+2i)c_1 e^{(1+2i)t}+(1-2i)c_2 e^{(1-2i)t}.$$

Queremos que

$$x(0) = c_1+c_2 = 1$$
$$y(0) = (1+2i)c_1+(1-2i)c_2 = 0.$$

Resolviendo el sistema obtenemos, $c_1 = \dfrac{2+i}{4}$, $c_2 = \dfrac{2-i}{4}$, y de aquí

$$x(t) = \frac{2+i}{4}e^{(1+2i)t} + \frac{2-i}{4}e^{(1-2i)t}$$
$$= e^t(\cos 2t - \tfrac{1}{2}\operatorname{sen} 2t)$$
$$y(t) = \dot{x}(t) = -\tfrac{5}{2}e^t \operatorname{sen} 2t.$$

Nótese que este sistema es equivalente a la ecuación de segundo orden

$$\ddot{x} - 2\dot{x} + 5x = 0$$

con las condiciones iniciales $x(0) = 1$, $\dot{x}(0) = 0$.

Problemas

1. Resuélvanse

a) $\dot{x}_1 = -3x_1$
$\dot{x}_2 = 2x_2$

b) $\dot{x}_1 = -2x_1$
$\dot{x}_2 = x_1 - x_2$

c) $\dot{x}_1 = -2x_1$
$\dot{x}_2 = x_1 - 2x_2$

d) $\dot{x} = 3x - 2y$
$\dot{y} = 3y$

e) $\dot{x} = 3x - 2y$
$\dot{y} = 2y, \quad x(0) = 1, \quad y(0) = 5$

f) $\dot{u} = 5u - 3v$
$\dot{v} = 12v, \quad u(0) = 0, \quad v(0) = 0$

g) $\dot{x}_1 = -3x_1$
$\dot{x}_2 = 2x_1 - 3x_2, \quad x_1(0) = a, \quad x_2(0) = b$

h) $\dot{x} = x$
$\dot{y} = x - y, \quad x(\sqrt{2}) = 5, \quad y(\sqrt{2}) = -7$

i) $\dot{x}_1 = \lambda_1 x_1$
$\dot{x}_2 = a_{21}x_1 + \lambda_2 x_2, \quad \lambda_1 \neq \lambda_2$

j) $\dot{x}_1 = \lambda_1 x_1$
$\dot{x}_2 = a_{21}x_1 + \lambda_2 x_2, \quad \lambda_1 = \lambda_2 = \lambda$

k) $\dot{x}_1 = \lambda_1 x_1 + a_{12}x_2$
$\dot{x}_2 = \lambda_2 x_2.$

2. Determínense los valores característicos y los vectores característicos correspondientes de las matrices:

a) $\begin{pmatrix} 2 & 0 \\ 1 & -3 \end{pmatrix}$ *b)* $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ *c)* $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ *d)* $\begin{pmatrix} \lambda_1 & 0 \\ 1 & \lambda_2 \end{pmatrix}$

e) $\begin{pmatrix} \lambda_0 & 0 \\ 1 & \lambda_0 \end{pmatrix}$ *f)* $\begin{pmatrix} -3 & 2 \\ 6 & -4 \end{pmatrix}$ g) $\begin{pmatrix} 1 & 0 \\ -c & -b \end{pmatrix}$ *h)* $\begin{pmatrix} 4 & -1 \\ 13 & -2 \end{pmatrix}$.

3. Una matriz $2\times 2A$ se dice que es *autoadjunta* (o hermitiana) si $a_{ij} = \bar{a}_{ji}$, $i, j = 1, 2$ ($\bar{a}_{ji}$ es el conjugado complejo de a_{ji}).

a) Pruébese que los valores característicos de una matriz autoadjunta son reales.

b) Si $\lambda_1 \neq \lambda_2$, pruébese que los vectores característicos correspondientes son ortogonales.

4. Resuélvanse los siguientes sistemas:

a) $\dot{x}_1 = x_1+3x_2$
$\quad\dot{x}_2 = 2x_2$

b) $\dot{x} = -x-y$
$\quad\dot{y} = x-5y$

c) $\dot{x} = x-y$
$\quad\dot{y} = 5x+3y$

d) $\dot{u} = 3u+6v$
$\quad\dot{v} = u+v$

e) $\dot{x} = 3x+2y$
$\quad\dot{y} = 3y$

f) $\dot{x} = 4x-3y$
$\quad\dot{y} = 3x-2y$

g) $\dot{x} = 2x-3y$
$\quad\dot{y} = 2x-2y, \qquad x(0) = 1, \qquad y(0) = 0$

h) $\dot{x} = y$
$\quad\dot{y} = -x-\frac{3}{2}y, \qquad x(0) = x_0, \qquad y(0) = y_0$

i) $\dot{x} = y$
$\quad\dot{y} = -a^2x-2ay, \qquad x(0) = x_0, \qquad y(0) = y_0$

j) $\dot{x} = 5x+3y-11$
$\quad\dot{y} = 8x+3y-14$.

Sugerencia: resuélvase el sistema $5x+3y-11=0$, $8x+3y-14=0$, y trasládese el origen al punto solución.

k) $\dot{x} = -2x+e^{-2t}$
$\quad\dot{y} = 3x-2y, \qquad x(0) = y(0) = 0.$

5. Cierta reacción entre dos sustancias se encuentra que obedece a la siguiente ley:

$$\dot{x} = -ax+by$$
$$\dot{y} = ax-by, \qquad a > 0,\ b > 0$$

donde $x(t)$ y $y(t)$ son las masas de las dos sustancias en el tiempo t.

a) Demuéstrese que la masa se conserva.

b) Determínense los productos finales de la reacción suponiendo que

$$x(0)+y(0) = M.$$

c) Cada solución $x(t)$, $y(t)$ define una curva γ en el plano XY. Proporciónese una representación geométrica de la reacción trazando tales curvas. Indíquese con flechas sobre las curvas cuál es la dirección de la reacción cuando el tiempo avanza. ¿Qué

es lo que tiene de significativo la recta que pasa por el origen en la dirección $\begin{pmatrix} b \\ a \end{pmatrix}$?

6. La probabilidad $p_n(t)$ de que en un contador se registren exactamente n partículas en el intervalo de tiempo $\langle 0, t \rangle$ viene dada por

$$p_0(t) = 1 - \alpha \int_0^t p_0(\tau)\, d\tau$$

$$p_1(t) = -\alpha \int_0^t [p_1(\tau) - p_0(\tau)]\, d\tau$$

$$\vdots$$

$$p_n(t) = -\alpha \int_0^t [p_n(\tau) - p_{n-1}(\tau)]\, d\tau$$

donde α es una constante positiva. Derívese una fórmula para $p_n(t)$ y pruébese que $\sum_{n=0}^{\infty} p_n(t) = 1$ para cualquier t.

7. Resuélvanse los siguientes sistemas:

a) $\dot{x} = y + z$

$\dot{y} = x + z$

$\dot{z} = x + y$

b) $\dfrac{dy}{dx} = 2z$

$\dfrac{dz}{dx} = 2w$

$\dfrac{dw}{dx} = 2y.$

8. Sean $\mathbf{x}^1$ y $\mathbf{x}^2$ soluciones de $\dot{\mathbf{x}} = A\mathbf{x}$ que satisfacen a $\mathbf{x}^1(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $\mathbf{x}^2(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. A éstas se les llama *soluciones principales*. Pruébese que la solución general de $\dot{\mathbf{x}} = A\mathbf{x}$ es $c_1\mathbf{x}^1 + c_2\mathbf{x}^2$.

9. Sean $\mathbf{x}^1$ y $\mathbf{x}^2$ un par de soluciones de $\dot{\mathbf{x}} = A\mathbf{x}$. Pruébese que:

a) Si $\mathbf{x}^1(0)$ y $\mathbf{x}^2(0)$ son linealmente independientes, entonces la solución general de $\dot{\mathbf{x}} = A\mathbf{x}$ es $c_1 x^1 + c_2 x^2$.

b) Si $\mathbf{x}^1(t_0)$ y $\mathbf{x}^2(t_0)$ son linealmente independientes para algún t_0, entonces $\mathbf{x}^1(t)$ y $\mathbf{x}^2(t)$ son linealmente independientes para todo t.

c) Si $\mathbf{x}^1(t_0)$ y $\mathbf{x}^2(t_0)$ son linealmente dependientes para algún t_0, entonces $\mathbf{x}^1(t)$ y $\mathbf{x}^2(t)$ son linealmente dependientes para todo t.

6. ECUACIONES DIFERENCIALES LINEALES DE SEGUNDO ORDEN CON COEFICIENTES CONSTANTES

La ecuación diferencial lineal de segundo orden con coeficientes constantes es

6.1 $$\ddot{x}+b\dot{x}+cx = f,$$

donde b y c son constantes dadas, y f es una función dada. Consideramos primero la ecuación homogénea

6.2 $$\ddot{x}+b\dot{x}+cx = 0.$$

Queremos encontrar todas las funciones x que satisfacen esta ecuación. Haciendo $\dot{x} = y$ podemos reducir inmediatamente 6.2 a un sistema de dos ecuaciones de primer orden:

6.3 $$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -cx-by. \end{aligned}$$

Este sistema lineal 6.3 es equivalente a 6.2 en el siguiente sentido: si x es una solución de 6.2 entonces $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}$ es una solución de 6.3, y si $\begin{pmatrix} x \\ y \end{pmatrix}$ es una solución de 6.3 entonces x es una solución de 6.2. Hay una sencilla correspondencia uno-uno entre las soluciones de las dos ecuaciones 6.2 y 6.3. Aquí

$$A = \begin{pmatrix} 0 & 1 \\ -c & -b \end{pmatrix}$$

y la ecuación característica de A es

$$\begin{vmatrix} -\lambda & 1 \\ -c & -b-\lambda \end{vmatrix} = \lambda^2+b\lambda+c = 0.$$

Las raíces características son

$$\lambda_1 = \tfrac{1}{2}(-b+\sqrt{b^2-4c})$$

y

$$\lambda_2 = \tfrac{1}{2}(-b-\sqrt{b^2-4})$$

De acuerdo con las ecuaciones 5.7′ y 5.9′, pág. 612, vemos que la solución general de 6.2 es

6.4 $$x(t) = c_1 e^{\lambda_1 t}+c_2 e^{\lambda_2 t} \quad \text{si} \quad \lambda_1 \neq \lambda_2$$

y

6.5 $$x(t) = (c_1 t+c_2)e^{\lambda t} \quad \text{si} \quad \lambda_1 = \lambda_2 = \lambda.$$

Estas son las soluciones generales para funciones complejas. Los números λ_1 y λ_2 lo mismo que los c_1 y c_2 pueden ser complejos. El interés usual generalmente está en las ecuaciones reales (b y c reales) y las soluciones reales. Aclaremos la cuestión de las soluciones reales para el sistema

6.6
$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 . \end{aligned}$$

Escrito como una ecuación vectorial simple, tenemos

6.7
$$\dot{\mathbf{x}} = A\mathbf{x},$$

donde $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Supongamos que la matriz A es real, con lo que queremos decir que los coeficientes a_{ij} son reales. En general, una solución x de 6.7 puede ser compleja, y podemos escribirla en la forma $\mathbf{x} = \mathbf{u} + i\mathbf{v}$ donde $\mathbf{u}$ y $\mathbf{v}$ son funciones vectoriales reales. Como A es real, $A\mathbf{u}$ y $A\mathbf{v}$ son también reales. De donde

$$\dot{\mathbf{x}} = \dot{\mathbf{u}} + i\dot{\mathbf{v}} = A(\mathbf{u} + i\mathbf{v}) = A\mathbf{u} + iA\mathbf{v},$$

e igualando las partes real e imaginaria:

$$\dot{\mathbf{u}} = A\dot{\mathbf{u}}, \qquad \dot{\mathbf{v}} = A\mathbf{v}.$$

Las partes real e imaginaria de una solución de 6.7 son también soluciones. Supongamos, para una solución cualquiera $x = \mathbf{u} + i\mathbf{v}$, que la condición inicial $\mathbf{x}(t_0) = \mathbf{x}^0$ es real. Entonces $\mathbf{v}(t_0) = \mathbf{0}$. Como $\dot{\mathbf{v}} = A\mathbf{v}$ con $\mathbf{v}(t_0) = 0$ tiene una solución única, se sigue que $\mathbf{v}$ es la solución trivial $\mathbf{0}$; es decir, $\mathbf{v}(t) = \mathbf{0}$ para toda t. Y hemos probado así que:

6.8 Lema. *Si la matriz A es real y si una solución de $\dot{\mathbf{x}} = A\mathbf{x}$ es real en algún instante t_0, entonces $\mathbf{x}(t)$ es real para todo t.*

El teorema de la unicidad y el anterior resultado pueden enunciarse como sigue para la ecuación diferencial lineal de segundo orden 6.2 que es equivalente al sistema 6.3.

6.9 Teorema. *Dados x_0, y_0 y t_0 hay una y sólo una solución de $\ddot{x} + b\dot{x} + cx = 0$ que satisface $x(t_0) = x_0$ y $\dot{x}(t_0) = y_0$. Si b, c, x_0 y y_0 son reales, entonces la solución de $\ddot{x} + b\dot{x} + cx = 0$ que satisface $x(t_0) = x_0$ y $\dot{x}(t_0) = y_0$ es real para todo t.*

En la hipótesis de que b y c son reales, entonces no es difícil probar que la solución real general de $\ddot{x} + b\dot{x} + cx = 0$ es como sigue:

Si λ_1 y λ_2 son reales, entonces

$$x(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}, \qquad \lambda_1 \neq \lambda_2$$

o bien

$$x(t) = (c_1 t + c_2)e^{\lambda_1 t}, \qquad \lambda_1 = \lambda_2,$$

donde c_1 *y* c_2 *son reales. Si* λ_1 *y* λ_2 *son complejos, entonces* $\lambda_2 = \bar{\lambda}_1$ *y*

$$x(t) = \text{Rl}\,(c_1 e^{\lambda_1 t})$$

donde c_1 *es un número complejo cualquiera.*

Así pues, en el caso de raíces complejas ($\lambda_1 = \beta + i\omega$, $\lambda_2 = \beta - i\omega$) la solución general puede expresarse en la forma

$$x(t) = e^{\beta t}(a_1 \cos \omega t + a_2 \operatorname{sen} \omega t)$$

o en la forma

$$x(t) = Ae^{\beta t} \cos(\omega t + \delta),$$

donde a_1, a_2, A y δ son números reales arbitrarios (problema 3).

6.10 Ejemplo. Obténganse todas las soluciones reales de

$\ddot{x} + \omega_0{}^2 x = 0,$ donde ω_0 es un número real distinto de cero.

Solución. La ecuación característica es

$$\lambda^2 + \omega_0{}^2 = 0$$

de forma que las raíces características son $\lambda_1 = i\omega_0$, $\lambda_2 = -i\omega_0$. La solución general es

$$x(t) = c_1 e^{i\omega_0 t} + c_2 e^{-i\omega_0 t}.$$

La solución real general es

$$x(t) = \text{Rl}\,(c_1 e^{i\omega_0 t}).$$

Si $c_1 = a_1 - ib_1$, entonces la solución real general es

$$x(t) = a_1 \cos \omega_0 t + b_1 \operatorname{sen} \omega_0 t$$

o bien

$$x(t) = a \cos(\omega_0 t + \delta),$$

donde $ae^{i\delta} = a_1 - ib_1$. Nótese, en particular, que $x(t) = \cos \omega_0 t$ es la solución que satisface $x(0) = 1$, $\dot{x}(0) = 0$, y $x(t) = \dfrac{1}{\omega_0} \operatorname{sen} \omega_0 t$ es la solución que satisface $x(0) = 0$, $\dot{x}(0) = 1$.

6.11 Ejemplo. Obténganse las soluciones de

$\ddot{x} - \omega_0{}^2 x = 0$, donde ω_0 es un número real distinto de cero,

que satisfacen: *a)* $x(0) = 1$, $\dot{x}(0) = 0$; *b)* $x(0) = 0$, $\dot{x}(0) = 1$.

SOLUCIÓN. La ecuación característica es

$$\lambda^2 - \omega_0{}^2 = 0,$$

de modo que $\lambda_1 = \omega_0$, $\lambda_2 = -\omega_0$. Por tanto, la solución general es

$$x(t) = c_1 e^{\omega_0 t} + c_2 e^{-\omega_0 t}.$$

a) $x(0) = c_1 + c_2 = 1$
$\dot{x}(0) = \omega_0 c_1 - \omega_0 c_2 = 0.$

Por tanto, $c_1 = c_2 = \frac{1}{2}$, y la solución que satisface esta condición inicial es

$$x(t) = \frac{e^{\omega_0 t} + e^{-\omega_0 t}}{2} = \cosh \omega_0 t.$$

b) Aquí

$$c_1 + c_2 = 0$$
$$\omega_0 c_1 - \omega_0 c_2 = 1,$$

y

$$c_1 = -c_2 = \frac{1}{2\omega_0}.$$

Así pues

$$x(t) = \frac{e^{\omega_0 t} - e^{-\omega_0 t}}{2\omega_0} = \frac{1}{\omega_0} \operatorname{senh} \omega_0 t.$$

Problemas

1. Resuélvanse

a) $\ddot{x} - x = 0$
b) $\ddot{x} + x = 0$
c) $\ddot{x} - 2\dot{x} + 3x = 0$
d) $9\ddot{x} + 9\dot{x} + 2x = 0$
e) $3\dfrac{d^2y}{dx^2} + \dfrac{dy}{dx} = 0$
f) $D^2x + 6Dx + 9x = 0$.

2. Resuélvanse

a) $\ddot{x} - 4x = 0, \quad x(0) = 0, \quad \dot{x}(0) = 10$
b) $\ddot{x} + 4x = 0, \quad x(0) = 0, \quad \dot{x}(0) = 10$
c) $\ddot{x} + 3\dot{x} = 0, \quad x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0$
d) $4\ddot{x} + 4\dot{x} + x = 0, \quad x(0) = 2, \quad \dot{x}(0) = -1$
e) $\ddot{x} + 4\dot{x} + 5x = 0, \quad x(0) = 2, \quad \dot{x}(0) = 0$
f) $4\ddot{x} + 12\dot{x} + 25x = 0, \quad x(0) = 5, \quad \dot{x}(0) = -3$.

3. Si b y c son reales y las raíces características de $\ddot{x} + b\dot{x} + cx = 0$ son

complejas, demuéstrese que las soluciones reales de esta ecuación diferencial son de la forma

a) $x(t) = c_1 e^{(\beta + i\omega)t} + \bar{c}_1 e^{(\beta - i\omega)t}$
$= 2e^{\beta t}\,\mathrm{Rl}\,(c_1 e^{i\omega t})$

b) $x(t) = e^{\beta t}(a_1 \cos \omega t + a_2 \operatorname{sen} \omega t)$

c) $x(t) = ae^{\beta t} \cos(\omega t + \delta)$.

(Exprésense β y ω en términos de b y c.)

4. Si $b > 0$ y $c > 0$, pruébese que los valores característicos de

$$\ddot{x} + b\dot{x} + cx = 0$$

tienen partes reales negativas. ¿Qué significa esto respecto al comportamiento de las soluciones cuando t tiende a infinito ?

5. Discútase la equivalencia entre las soluciones reales del sistema

$$\dot{x} = y$$
$$\dot{y} = -x$$

y las soluciones de

$$\dot{z} + iz = 0.$$

Sea $z = x + iy$. Obténganse de aquí las soluciones reales del sistema.

6. Generalícese el problema 5 comenzando con

$$\dot{z} + (a + ib)z = 0$$

7. Usando la sustitución $y(\tau) = x(e^\tau)$ redúzcase el problema de resolver (la ecuación equidimensional)

$$\ddot{x}(t) + \frac{b}{t}\dot{x}(t) + \frac{c}{t^2}x(t) = 0 \quad \text{sobre } \langle 0, \infty \rangle$$

al de resolver la ecuación lineal homogénea

$$\ddot{y} + (b-1)\dot{y} + cy = 0.$$

8. Resuélvase

a) $t^2\ddot{x}(t) - 2t\dot{x}(t) + 2x(t) = 0$ sobre $\langle 0, \infty \rangle$

b) $\ddot{x}(t) + \frac{1}{t}\dot{x}(t) - \frac{n^2}{t^2}x(t) = 0$ sobre $\langle 0, \infty \rangle$, $n \neq 0$

c) $t^2\ddot{x}(t) + t\dot{x}(t) = 0$ sobre $\langle 0, \infty \rangle$.

9. Resuélvase

a) $2\ddot{x} - \dot{x} - 3x = 0$

b) $L\ddot{q}+R\dot{q}+\frac{1}{C}q=0$ dados

$L=100\times 10^{-6}$, $R=30\times 10^{2}$, $C=100\times 10^{-12}$,
$q(0)=100\times 10^{-10}$, $\dot{q}(0)=0$

c) Resuélvase *b* cuando $R=0$.

7. LA ECUACIÓN COMPLETA $\dot{\mathbf{x}}=A\mathbf{x}+\mathbf{f}$

El sistema de ecuaciones diferenciales

7.1
$$\begin{aligned}\dot{x}_1 &= a_{11}x_1+a_{12}x_2+f_1\\ \dot{x}_2 &= a_{21}x_1+a_{22}x_2+f_2\end{aligned}$$

se llama sistema "completo" en contraste con el sistema

7.2
$$\begin{aligned}\dot{x}_1 &= a_{11}x_1+a_{12}x_2\\ \dot{x}_2 &= a_{21}x_1+a_{22}x_2\end{aligned}$$

que se llama sistema "homogéneo". Con

$$\mathbf{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix},\quad A=(a_{ij})=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix},\quad \mathbf{f}=\begin{pmatrix}f_1\\f_2\end{pmatrix},$$

estos sistemas pueden escribirse

7.3
$$\dot{\mathbf{x}}=A\mathbf{x}+\mathbf{f}$$

y

7.4
$$\dot{\mathbf{x}}=A\mathbf{x}.$$

Suponemos que **f** es continua sobre $\langle-\infty,\infty\rangle$. Será claro, sin embargo, que la discusión puede restringirse a cualquier intervalo $\mathscr{I}$ en el que **f** sea continua.

La introducción de la notación vector-matricial nos permitirá ver que la teoría del sistema lineal 7.1 no es más difícil que la de la ecuación simple

$$\dot{x}=ax+f.$$

Aquí e^{-at} es un factor integrante:

$$e^{-at}(\dot{x}(t)-ax(t))=\frac{d}{dt}(e^{-at}x(t))=e^{-at}f(t);$$

y

$$x(t)=ce^{at}+e^{at}\int_0^t e^{-as}f(s)\,ds$$

es la solución general de la ecuación completa.

Definiremos ahora la exponencial matricial e^{-At} y demostraremos que es un factor de integración de 7.3.

Una matriz se dice que es *no singular* si $A\mathbf{x} = \mathbf{0}$ implica $\mathbf{x} = \mathbf{0}$. Denotemos el determinante de A por det (A), es decir

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

En esta notación la ecuación característica de A es

$$\varphi(\lambda) = \det(A-\lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = 0, \text{ donde } I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Una matriz A se dice que tiene una inversa si existe una matriz A^{-1} con la propiedad de que

$$A^{-1}A = AA^{-1} = I.$$

A^{-1} se llama el *inverso* de A. Es fácil demostrar que A^{-1} es único. Supongamos que A es también un inverso. Entonces

$$A^* = A^*(AA^{-1}) = (A^*A)A^{-1} = A^{-1}.$$

Si A tiene una matriz inversa entonces

$$A\mathbf{x} = \mathbf{0} \Rightarrow A^{-1}(A\mathbf{x}) = \mathbf{0}$$
$$\Rightarrow (A^{-1}A)\mathbf{x} = I\mathbf{x} = \mathbf{x} = \mathbf{0}.$$

Así pues, si A tiene una inversa entonces A es no singular. También puede probarse lo recíproco. Por lo que sabemos acerca de la solución de ecuaciones algebraicas lineales vemos que *una matriz A es no singular* ($A\mathbf{x} = \mathbf{0}$ sólo si $\mathbf{x} = \mathbf{0}$) *si y sólo si* det $(A) \neq 0$. Así pues, si A es no singular, entonces dada $\mathbf{y}$ la ecuación

$$\mathbf{y} = A\mathbf{x}$$

tiene una solución única

$$\mathbf{x} = \frac{1}{\det(A)}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}\mathbf{y} = B\mathbf{y}.$$

Es fácil comprobar que $AB = BA = I$ y, por tanto, que $B = A^{-1}$. Esto completa la prueba del siguiente lema.

7.5 Lema. *Una matriz A es no singular si y sólo si tiene una inversa.*

Sea $X = (x_{ij})$ una función matricial (con valores matrices) definida por

$$X(t) = (x_{ij}(t)) = \begin{pmatrix} x_{11}(t) & x_{12}(t) \\ x_{21}(t) & x_{22}(t) \end{pmatrix}.$$

Los conceptos y operaciones sobre funciones matriciales son los mismos que los correspondientes sobre funciones vectoriales. Consideraremos la matriz como un vector tetradimensional: $X(t)$ es continua sobre un intervalo $\mathscr{I}$ si todos los x_{ij} son continuos sobre $\mathscr{I}$,

$$\dot{X}(t) = (\dot{x}_{ij}(t)),$$

e

$$\int_a^b X(t)\,dt = \left(\int_a^b x_{ij}(t)\,dt\right).$$

Por ejemplo

$$\frac{d}{dt}\begin{pmatrix} \cos t & \operatorname{sen} t \\ -\operatorname{sen} t & \cos t \end{pmatrix} = \begin{pmatrix} -\operatorname{sen} t & \cos t \\ -\cos t & -\operatorname{sen} t \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \cos t & \operatorname{sen} t \\ -\operatorname{sen} t & \cos t \end{pmatrix}.$$

De donde $X(t) = \begin{pmatrix} \cos t & \operatorname{sen} t \\ -\operatorname{sen} t & \cos t \end{pmatrix}$ es solución de $\dot{X} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} X$.

Se prueba entonces fácilmente (problema 1) que

7.6 $$\frac{d}{dt}(XY) = \dot{X}Y + X\dot{Y}$$

7.7 $$\frac{d}{dt}(X\mathbf{v}) = \dot{X}\mathbf{v} + X\dot{\mathbf{v}}.$$

Consideremos ahora la ecuación diferencial matricial

7.8 $$\dot{X} = AX,$$

donde A es una matriz constante. Como

$$AX = \begin{pmatrix} a_{11}x_{11} + a_{12}x_{21} & a_{11}x_{12} + a_{12}x_{22} \\ a_{21}x_{11} + a_{22}x_{21} & a_{21}x_{12} + a_{22}x_{22} \end{pmatrix},$$

vemos que X es una solución de 7.8 si y sólo si

$$\begin{pmatrix} \dot{x}_{11} \\ \dot{x}_{21} \end{pmatrix} = A\begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix}, \qquad \begin{pmatrix} \dot{x}_{12} \\ \dot{x}_{22} \end{pmatrix} = A\begin{pmatrix} x_{12} \\ x_{22} \end{pmatrix};$$

es decir, si y sólo si los vectores columna de X son soluciones de la ecuación diferencial vectorial

7.9 $$\dot{\mathbf{x}} = A\mathbf{x}.$$

Esta es exactamente nuestra razón para introducir la ecuación matricial. Es un lenguaje conveniente (que puede generalizarse para n dimensiones) para hablar de un par de soluciones de 7.9 (o, en general, de n soluciones de un sistema de n ecuaciones diferenciales lineales). Las dos soluciones en que estamos interesados son las soluciones de 7.9 que satisfacen $\mathbf{x}(0) = \begin{pmatrix}1\\0\end{pmatrix}$ y $\mathbf{x}(0) = \begin{pmatrix}0\\1\end{pmatrix}$. Se llaman a éstas las soluciones *principales* de 7.9. La *solución matricial principal* de 7.8 es la solución de 7.8 que satisface $X(0) = I$. Denotemos por P a la solución matricial principal de 7.8. Las columnas de P son las soluciones principales de 7.9.

Como las soluciones matriciales de 7.8 se corresponden con pares de soluciones de 7.9, sabemos, según el teorema 5.12 (pág. 614) que: *dada una matriz constante X_0 hay una matriz solución y sólo una matriz solución de* 7.8 *que satisface a* $X(0) = X_0$. Usamos ahora este resultado de existencia y unicidad para demostrar que la solución matricial principal P tiene las propiedades de una exponencial.

Definamos la función matricial Y por $Y(t) = P(t+\tau)$, donde τ es un número real cualquiera. Entonces

$$\dot{Y}(t) = \frac{d}{dt}[P(t+\tau)] = \dot{P}(t+\tau) = AP(t+\tau) = AY(t),$$

lo que nos dice que $Y(t)$ es la solución de 10.8 que satisface $Y(0) = P(\tau)$. Ahora bien, $Z(t) = P(t)\,P(\tau)$ es también una solución de 7.8 que satisface la misma condición inicial: $Z(0) = P(0)\,P(\tau) = IP(\tau) = P(\tau)$. De donde $Z = Y$; es decir

7.10 $$P(t)\,P(\tau) = P(t+\tau)$$

para toda t y toda τ. Esta es la ley de los exponentes y, como $\dot{P}(t) = AP(t)$ y $P(0) = I$, la solución matricial principal se escribe a menudo como una exponencial:

$$P(t) = e^{At},$$

y, por analogía con la exponencial real y compleja, se llama *exponencial matricial.* Podemos entonces escribir 7.10 en la forma

$$e^{At}e^{A\tau} = e^{A(t+\tau)}.$$

Nótese entonces que

$$e^{At}e^{-At} = e^{-At}e^{At} = e^0 = P(0) = I.$$

De donde $P(t) = e^{At}$ es no singular para toda t y $P^{-1}(t) = e^{-At}$.

Con esta matriz exponencial podemos fácilmente resolver la ecuación completa

7.11 $$\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}.$$

Como $\frac{d}{dt}(e^{-At}) = \frac{d}{dt}P(-t) = -\dot{P}(-t) = -AP(-t) = -Ae^{-At}$, podemos imaginarnos, basándonos en lo que sucede en el caso unidimensional, que e^{-At} es un factor de integración de 7.11. Para comprobarlo, necesitamos saber que A y e^{-At} conmutan; es decir, que $Ae^{-At} = e^{-At}A$. Sean

$$Y(t) = Ae^{-At}, \quad Z(t) = e^{-At}A.$$

Entonces

$$\dot{Y}(t) = -A^2e^{-At} = -AY(t)$$

y

$$\dot{Z}(t) = -Ae^{-At}A = -AZ(t).$$

Como $Z(0) = Y(0) = A$, usando de nuevo el argumento de la unicidad, tenemos $Z = Y$. De donde $Ae^{-At} = e^{-At}A$. Por tanto,

$$\frac{d}{dt}(e^{-At}\mathbf{x}(t)) = e^{-At}\dot{\mathbf{x}}(t) - Ae^{-At}\mathbf{x}(t) = e^{-At}\dot{\mathbf{x}}(t) - e^{-At}A\mathbf{x}(t)$$

$$= e^{-At}(\dot{\mathbf{x}}(t) - A\mathbf{x}(t)).$$

Supongamos que $\mathbf{x}$ es una solución de la ecuación completa 7.11 que satisface $\mathbf{x}(0) = \mathbf{c}$. Entonces

$$e^{-At}(\dot{\mathbf{x}}(t) - A\mathbf{x}(t)) = e^{-At}\mathbf{f}(t)$$

o bien

$$\frac{d}{dt}(e^{-At}\mathbf{x}(t)) = e^{-At}\mathbf{f}(t).$$

Por tanto

$$e^{-At}\mathbf{x}(t) - \mathbf{x}(0) = \int_0^t e^{-As}\mathbf{f}(s)\,ds$$

y

7.12 $$\mathbf{x}(t) = e^{At}\mathbf{c} + \int_0^t e^{A(t-s)}\mathbf{f}(s)\,ds.$$

Es fácil comprobar que $\mathbf{x}(t)$ es una solución de 7.11 y que $\mathbf{x}(0) = \mathbf{c}$. De donde 7.12 es la solución de la ecuación completa 7.11 que satisface $\mathbf{x}(0) = \mathbf{c}$; es la solución general de 7.11. El término $e^{At}\mathbf{c}$ es la solución general de la ecuación homogénea y $e^{At}\int_0^t e^{-As}\mathbf{f}(s)\,ds$ es una solución

particular de la ecuación completa [es la solución de 7.11 que satisface $\mathbf{x}(0)=\mathbf{0}$, problema 5].

7.13 Ejemplo. Resuélvase

$$\ddot{x}+\omega_0{}^2 x = \operatorname{sen}\omega t, \quad \omega \neq 0,\ \omega_0 \neq 0.$$

SOLUCIÓN. Un sistema equivalente es

$$\begin{aligned}\dot{x} &= y\\ \dot{y} &= -\omega_0{}^2 x + \operatorname{sen}\omega t.\end{aligned}$$

Aquí $A = \begin{pmatrix} 0 & 1\\ -\omega_0{}^2 & 0\end{pmatrix}$, $\mathbf{f}(t) = \begin{pmatrix} 0\\ \operatorname{sen}\omega t\end{pmatrix}$. Las soluciones principales de la ecuación reducida son (véase ejemplo 6.10, pág. 621)

$$\mathbf{x}^1 = \begin{pmatrix} \cos\omega_0 t\\ -\omega_0 \operatorname{sen}\omega_0 t\end{pmatrix}, \qquad \mathbf{x}^2 = \begin{pmatrix} \dfrac{1}{\omega_0}\operatorname{sen}\omega_0 t\\ \\ \cos\omega_0 t\end{pmatrix}.$$

La solución matricial principal es

$$P(t) = e^{At} = \begin{pmatrix} \cos\omega_0 t & \dfrac{1}{\omega_0}\operatorname{sen}\omega_0 t\\ \\ -\omega_0\operatorname{sen}\omega_0 t & \cos\omega_0 t\end{pmatrix}.$$

Usando 7.12 tenemos que

$$\begin{pmatrix} x(t)\\ y(t)\end{pmatrix} = \begin{pmatrix} \cos\omega_0 t & \dfrac{1}{\omega_0}\operatorname{sen}\omega_0 t\\ \\ -\omega\operatorname{sen}\omega_0 t & \cos\omega_0 t\end{pmatrix}\begin{pmatrix} c_1\\ c_2\end{pmatrix}$$

$$+\int_0^t \begin{pmatrix} \cos\omega_0(t-s) & \dfrac{1}{\omega_0}\operatorname{sen}\omega_0(t-s)\\ \\ -\omega_0\operatorname{sen}\omega_0(t-s) & \cos\omega_0(t-s)\end{pmatrix}\begin{pmatrix} 0\\ \operatorname{sen}\omega s\end{pmatrix} ds.$$

Como estamos solamente interesados en x, tenemos

$$x(t) = c_1\cos\omega_0 t + \frac{1}{\omega_0}c_2\operatorname{sen}\omega_0 t + \frac{1}{\omega_0}\int_0^t \operatorname{sen}\omega_0(t-s)\operatorname{sen}\omega s\, ds.$$

Por tanto, si $\omega \neq \omega_0$,

$$x(t) = c_1 \cos \omega_0 t + \frac{1}{\omega_0} c_2 \operatorname{sen} \omega_0 t - \frac{\omega}{\omega_0(\omega_0^2 - \omega^2)} \operatorname{sen} \omega_0 t + \frac{1}{\omega_0^2 - \omega^2} \operatorname{sen} \omega t$$

$$= k_1 \cos \omega_0 t + k_2 \operatorname{sen} \omega_0 t + \frac{1}{\omega_0^2 - \omega^2} \operatorname{sen} \omega t.$$

Si $\omega = \omega_0$,

$$x(t) = k_1 \cos \omega_0 t + k_2 \operatorname{sen} \omega_0 t - \frac{1}{2\omega_0} t \cos \omega_0 t.$$

Si $\omega = \omega_0$, tenemos lo que se llama *resonancia*. Si $\omega \neq \omega_0$, las oscilaciones son acotadas para todo t y ω fijo, aunque la amplitud del término sen ωt tiende a ∞ cuando $\omega \to \omega_0$. En el caso de resonancia $\omega = \omega_0$, $\lim_{t \to \infty} x(t)$ no existe. Este fenómeno de resonancia se discute con más detalle en la sección 11.

Problemas

1. Pruébense: *a*) 7.6 *b*) 7.7

2. Pruébese que: $\frac{d}{dt} X^{-1} = -X^{-1}\left(\frac{dX}{dt}\right) X^{-1}$.

3. Pruébese que

a) $e^{\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} t} = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix}$ *b*) $e^{\begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix} t} = e^{\lambda t} \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$

c) $e^{\begin{pmatrix} \lambda_1 & 0 \\ 1 & \lambda_2 \end{pmatrix} t} = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ \dfrac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\lambda_1 - \lambda_2} & e^{\lambda_2 t} \end{pmatrix}$ *d*) $e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} t} = \begin{pmatrix} \cos t & \operatorname{sen} t \\ -\operatorname{sen} t & \cos t \end{pmatrix}$

e) $e^{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} t} = \begin{pmatrix} \cosh t & \operatorname{senh} t \\ \operatorname{senh} t & \cosh t \end{pmatrix}$.

4. Resuélvase:

a) $\dot{x} = y$
$\dot{y} = -x - 2y + e^t, \quad x(0) = y(0) = 0$

b) $\dot{x} = y$
$\dot{y} = x + te^{-t}, \quad x(0) = 1, \quad y(0) = 0$

$c)$ $\dot{x} = x-y+t$
$\dot{y} = -x+y+t^2, \quad x(0)=1, \quad y(0)=3$

$d)$ $\dot{x} = ax-by$
$\dot{y} = bx+ay+e^{-at}, \quad x(0)=x_0, \quad y(0)=y_0.$

5. Pruébese que

$$\mathbf{x}(t) = \int_0^t e^{A(t-s)}\mathbf{f}(s)\,ds$$

es la solución de $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$ que satisface $\mathbf{x}(0)=\mathbf{0}$.

6. Pruébese que

$$e^{At} = \sum_{n=0}^{\infty} \frac{1}{n!}A^n t^n.$$

8. LA ECUACIÓN COMPLETA $\ddot{x}+b\dot{x}+cx=f$

La ecuación diferencial lineal de segundo orden

8.1 $$\ddot{x}+b\dot{x}+cx = f$$

es, desde luego, un caso particular del sistema que estudiamos en la sección 7. Pero es un caso particularmente importante, y la solución general tiene una forma sencilla. Un sistema equivalente es:

8.2 $$\begin{aligned}\dot{x} &= y\\ \dot{y} &= -cx-by+f.\end{aligned}$$

Suponemos que f es continua sobre $\langle -\infty, \infty\rangle$, o bien podíamos suponer que f es continua sobre $\mathscr{I}$ y restringir la discusión a $\mathscr{I}$. Sean

$$A = \begin{pmatrix} 0 & 1\\ -c & -b\end{pmatrix} \quad \text{y} \quad \mathbf{f} = \begin{pmatrix} 0\\ f\end{pmatrix}.$$

Sabemos cómo obtener las soluciones principales $\mathbf{p}^1 = \begin{pmatrix} p_{11}\\ p_{21}\end{pmatrix}$ y $\mathbf{p}^2 = \begin{pmatrix} p_{12}\\ p_{22}\end{pmatrix}$ del sistema homogéneo

8.3 $$\begin{aligned}\dot{x} &= y\\ \dot{y} &= -cx-by.\end{aligned}$$

La función p_{11} es la solución de la ecuación homogénea

8.4 $$\ddot{x}+b\dot{x}+cx = 0$$

que satisface $x(0)=1$ y $\dot{x}(0)=0$. Análogamente p_{12} es la solución de 8.4 que satisface $x(0)=0$, $\dot{x}(0)=1$. Sabemos, por tanto, que la solución

matricial principal $P(t) = (p_{ij}(t)) = e^{At}$ y, por la ecuación 7.12, la solución general de la ecuación completa 8.2 es

8.5
$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{At}\mathbf{c} + \int_0^t e^{A(t-s)}\mathbf{f}(s)\,ds.$$

Aquí

$$e^{A(t-s)}\mathbf{f}(s) = \begin{pmatrix} p_{11}(t-s) & p_{12}(t-s) \\ p_{21}(t-s) & p_{22}(t-s) \end{pmatrix} \begin{pmatrix} 0 \\ f(s) \end{pmatrix}.$$

Nuestro interés está en $x(t)$ y la solución general de la ecuación completa 8.1 es

8.6
$$x(t) = c_1 p_{11}(t) + c_2 p_{12}(t) + \int_0^t p_{12}(t-s) f(s)\,ds.$$

Haciendo $v(t) = p_{11}(t)$ y $w(t) = p_{12}(t)$, tenemos

8.7 Teorema. *Sean v y w las soluciones de $\ddot{x} + b\dot{x} + cx = 0$ que satisfacen $v(0) = 1$, $\dot{v}(0) = 0$ y $w(0) = 0$, $\dot{w}(0) = 1$. Entonces la solución general de $\ddot{x} + b\dot{x} + cx = f$ es*

$$x(t) = c_1 v(t) + c_2 w(t) + \int_0^t w(t-s) f(s)\,ds.$$

La integral $\int_0^t w(t-s) f(s)\,ds$ es la solución particular de $\ddot{x} + b\dot{x} + cx = f$ que satisface $x(0) = \dot{x}(0) = 0$. (Véase en conexión con esto el problema 3, pág. 635.) La diferencia entre dos soluciones cualesquiera de la ecuación completa es una solución de las ecuaciones homogéneas. Podemos por ello enunciar: *la solución general de la ecuación completa es la solución general de la ecuación homogénea más una solución particular cualquiera de la ecuación completa.*

8.8 Ejemplo. Resuélvase $\ddot{x} + 2\dot{x} + x = e^t$, $x(0) = \dot{x}(0) = 0$.

SOLUCIÓN 1. La ecuación característica es

$$\lambda^2 + 2\lambda + 1 = (\lambda + 1)^2 = 0$$

y -1 es una raíz múltiple. De donde la solución general de la ecuación homogénea es $(c_1 + c_2 t)e^{-t}$, y vemos que $w(t) = te^{-t}$. Por tanto,

$$\begin{aligned} x(t) &= \int_0^t (t-s)e^{-(t-s)} e^s\,ds \\ &= te^{-t}\int_0^t e^{2s}\,ds - e^{-t}\int_0^t se^{2s}\,ds \\ &= \tfrac{1}{4}e^t - \tfrac{1}{4}(1+2t)e^{-t}, \end{aligned}$$

que es la solución buscada.

SOLUCIÓN 2. Busquemos una solución particular de la forma ce^t. Entonces, sustituyendo en la ecuación diferencial, tenemos como ecuación para c

$$c+2c+c = 1 \quad \text{o} \quad c = \tfrac{1}{4}.$$

De donde la solución general es

$$x(t) = (c_1+c_2t)e^{-t}+\tfrac{1}{4}e^t.$$

Las condiciones iniciales son

$$x(0) = c_1+\tfrac{1}{4} = 0$$
$$\dot{x}(0) = -c_1+c_2+\tfrac{1}{4} = 0.$$

Por tanto, $c_1 = -\frac{1}{4}$, $c_2 = -\frac{1}{2}$, y la solución buscada es

$$x(t) = -\tfrac{1}{4}(1+2t)e^{-t}+\tfrac{1}{4}e^t.$$

8.9 Ejemplo. Resuélvase $\ddot{x}+x = f$, $x(0) = \dot{x}(0) = 0$, donde

$$f(t) = \begin{cases} 0, & t<0 \\ t, & 0 \leqslant t < \pi \\ 2\pi - t, & \pi \leqslant t < 2\pi \\ 0, & 2\pi \leqslant t. \end{cases}$$

SOLUCIÓN. La solución w de la ecuación homogénea que satisface $w(0) = 0$, $\dot{w}(0) = 1$ es

$$w(t) = \operatorname{sen} t.$$

De donde

$$x(t) = \int_0^t f(s) \operatorname{sen}(t-s)\, ds.$$

Por tanto

$$x(t) = \begin{cases} 0, \quad t<0 \\ \displaystyle\int_0^t s \operatorname{sen}(t-s)\, ds, \quad 0 \leqslant t < \pi \\ \displaystyle\int_0^\pi s \operatorname{sen}(t-s)\, ds + \int_\pi^t (2\pi - s) \operatorname{sen}(t-s)\, ds, \quad \pi \leqslant t < 2\pi \\ \displaystyle\int_0^\pi s \operatorname{sen}(t-s)\, ds + \int_\pi^{2\pi} (2\pi - s) \operatorname{sen}(t-s)\, ds, \quad 2\pi \leqslant t. \end{cases}$$

$$= \begin{cases} 0, & t < 0 \\ t - \operatorname{sen} t, & 0 \leqslant t < \pi \\ 2\pi - t - 3 \operatorname{sen} t, & \pi \leqslant t < 2\pi \\ -4 \operatorname{sen} t, & 2\pi \leqslant t. \end{cases}$$

La función f, que en una aplicación, puede ser una fuerza o voltaje aplicado, puede considerarse como una energía de cierto tipo que entra en un sistema (figura 3). En el instante $t = 0$ el sistema está parado en su posición de equilibrio; y la solución x puede verse como la energía de salida (figura 4) que corresponde a la de entrada f.

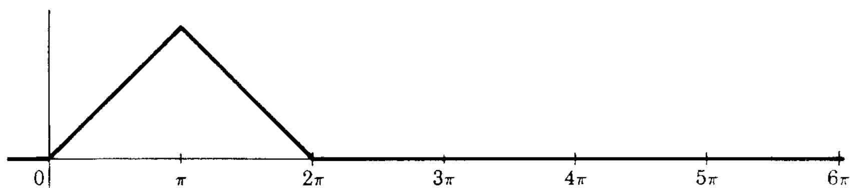


FIGURA 3

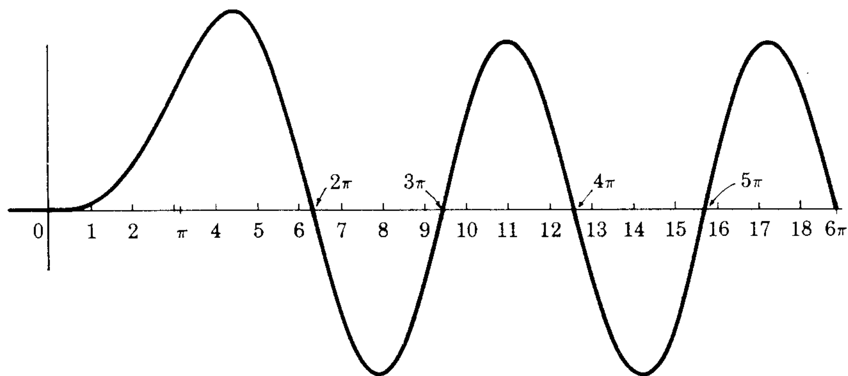


FIGURA 4

Problemas

1. Resuélvanse:

a) $\ddot{x} + 3\dot{x} + 2x = t^2 - t, \quad x(0) = \dot{x}(0) = 0$
b) $\ddot{x} + 2\dot{x} + x = \cos t, \quad x(0) = \dot{x}(0) = 0$
c) $2\ddot{x} + 7\dot{x} + 3x = 5 - e^t, \quad x(0) = 0, \quad \dot{x}(0) = 1$
d) $\ddot{x} + 4\dot{x} = e^t, \quad x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0$

e) $\ddot{x}+c^2x = e^t, \quad x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0$

f) $\dfrac{d^2y}{dx^2} - \dfrac{dy}{dx} - 2y = 3e^x, \quad y(1) = y'(1) = 0$

g) $\ddot{x}+2\dot{x}+2x = f, \quad x(0) = \dot{x}(0) = 0,$

donde $f(t) = \begin{cases} 0, & t < 0 \\ t, & 0 \leqslant t < 1 \\ 1, & 1 \leqslant t\,. \end{cases}$

2. Pruébese que para el caso $\omega \neq \omega_0$ (no resonancia), la ecuación diferencial

$$\ddot{x}+\omega_0{}^2x = a\cos\omega t$$

tiene una solución periódica de periodo $\dfrac{2\pi}{\omega}$. Trácese la amplitud (el valor máximo) de esta solución periódica como función de ω.

3. Una función se dice que es *continua a trozos* sobre un intervalo $[a, b]$ si f es continua en todos, salvo un número finito de puntos, y en cada punto de discontinuidad existen límites a la derecha y a la izquierda. Decimos que f es continua a trozos en $\langle -\infty, \infty\rangle$ si f es continua a trozos sobre todo intervalo $[a, b]$. Cuando f es continua a trozos sobre $\langle -\infty, \infty\rangle$, decimos que y es una solución de $(*)$ $\ddot{x}+b\dot{x}+cx=f$ si y es continua sobre $\langle -\infty, \infty\rangle$ y satisface $(*)$ en todo punto donde f es continua. Bajo la hipótesis de que f es continua a trozos sobre $\langle -\infty, \infty\rangle$ pruébese que $x(t) = c_1v(t)+c_2w(t) + \displaystyle\int_0^t w(t-s)f(s)\,ds$ es la solución general de $(*)$ sobre $\langle -\infty, \infty\rangle$; v y w definidos como en el teorema 8.7.

Nota. Esta es la extensión a impulsos de entrada continuos a trozos f del teorema 8.7. Si g es continua a trozos sobre $[a, b]$, entonces g es integrable sobre $[a, b]$. Si g es integrable sobre $[a, b]$ entonces $G(t) = \displaystyle\int_a^t g$ es continua sobre $[a, b]$ y en cada punto t de $[a, b]$ donde g es continua $G'(t)=g(t)$.

4. Resuélvase

$$\ddot{x}+x = f, \quad x(0) = \dot{x}(0) = 0$$

donde

a) $f(t) = \begin{cases} 0, & t < 0 \\ 1, & 0 \leqslant t < 1 \\ 0, & t \geqslant 1 \end{cases}$ *b*) $f(t) = \begin{cases} 0, & t < 0 \\ t, & 0 \leqslant t < 1 \\ 0, & t \geqslant 1 \end{cases}$

c) $f(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geqslant 0. \end{cases}$

Trácese la gráfica de los impulsos de entrada f y de las soluciones x (las "salidas").

5. Supongamos que f y g son continuas sobre un intervalo $\mathscr{I}$ con $0 \in \mathscr{I}$. La función $f * g$ definida por

$$(f * g)(t) = \int_0^t f(t-s)\, g(s)\, ds$$

se llama *convolución* de f y g. Pruébese que $f * g = g * f$; es decir,

$$\int_0^t f(t-s)\, g(s)\, ds = \int_0^t f(s)\, g(t-s)\, ds.$$

Nota. Este resultado puede a veces simplificar el cálculo de $\int_0^t w(t-s)\, f(s)\, ds$. La convolución puede también definirse por

$$(f * g)(t) = \int_{-\infty}^{\infty} f(t-s)\, g(s)\, ds.$$

Las dos definiciones son equivalentes si $f(t) = g(t) = 0$ para $t < 0$.

6. Resuélvase

$$\ddot{x} + 3\dot{x} + 2x = f, \quad x(0) = \dot{x}(0) = 0,$$

donde f es

a) la f del problema 4*a*
b) la f del problema 4*b*
c) la f del problema 4*c*.

7. La *función u del escalón unitario* es la función definida por $u(t) = 0$, $t < 0$, y $u(t) = 1$, $t \geqslant 0$. Sea y la solución de $x + bx + cx = u$ que satisface a $x(0) = \dot{x}(0) = 0$. (a y se le llama *la respuesta del sistema a la función del escalón unitario*)

a) Pruébese que sobre $[0, \infty]$ $\dot{y} = w$ y, por tanto, $y(t) = \int_0^t w(s)\, ds$, donde w es la función dada en el teorema 8.7. Así pues, se sabe o puede medirse la respuesta de este sistema lineal a una función de escalón unitario, entonces puede calcularse la respuesta $x(t) = \int_0^t w(t-s)\, f(s)\, ds$ a cualquier impulso de entrada f.

b) La respuesta de un sistema a una función de escalón unitario

es (para $t \geqslant 0$) $1-(1+t)e^{-t}$. Determínese la respuesta de este sistema lineal al pulso cuadrado

$$f(t) = \begin{cases} 0, & t < 0 \\ 1, & 0 \leqslant t < \delta \\ 0, & t \geqslant \delta. \end{cases}$$

9. EL PRINCIPIO DE SUPERPOSICIÓN

El principio de superposición, que en esta sección probaremos y discutiremos, es la propiedad característica de los sistemas lineales. Es la falta de este principio la que complica el estudio de los sistemas no lineales. Hasta cierto punto, este principio ha llevado a un estudio de los sistemas lineales completamente separado de las ecuaciones diferenciales, lo que ha probado ser causa de confusión. Los métodos lineales son adecuados dentro de su rango de aplicabilidad, pero no sirven en general más que para sistemas lineales y no se debe esperar poder, en general, aplicar métodos lineales al estudio de sistemas no lineales. No queremos asegurar con esto que el estudio de los sistemas lineales no es importante. Las rectas son el punto de partida más natural para el estudio de la geometría, las ecuaciones lineales son el punto natural de partida para el estudio de las ecuaciones algebraicas y trascendentes y, de igual modo, lo son las ecuaciones diferenciales lineales para el estudio de las ecuaciones diferenciales. Pero es, sin embargo, de esperar que la no linealidad planteará nuevos problemas, requerirá métodos diferentes y que los sistemas no lineales exhibirán comportamientos que no podrán encontrarse en los sistemas lineales. El comportamiento no lineal no es forzosamente indeseable. En realidad, lo que quiere decirse es que en los sistemas no lineales pueden esperarse que ocurran cosas que en los sistemas lineales nunca ocurrirá. Significa también que para muchos propósitos la aproximación lineal de un sistema no puede por sí misma ser adecuada para la explicación de ciertos fenómenos.

Consideremos el sistema lineal

9.1 $$\dot{\mathbf{x}}(t) = A(t)\,\mathbf{x}(t)+\mathbf{f}(t),$$

donde suponemos que la función matricial A y la función vectorial $\mathbf{f}$ son continuas sobre $\langle -\infty, \infty\rangle$. El principio de superposición es, como veremos, una consecuencia directa de la linealidad de $A(t)\,\mathbf{x}(t)$: $A(t)\,[c_1\mathbf{x}^1(t)+c_2\mathbf{x}^2(t)] = c_1 A(t)\,\mathbf{x}^1(t)+c_2 A(t)\,\mathbf{x}^2(t)$.

9.2 Teorema. (El principio de superposición.) *Si* $\mathbf{y}^1$ *es una solución de* $\dot{\mathbf{x}}(t) = A(t)\,\mathbf{x}(t)+\mathbf{f}^1(t)$ *y* $\mathbf{y}^2$ *es una solución de* $\dot{\mathbf{x}}(t) = A(t)\,\mathbf{x}(t)+\mathbf{f}^2(t)$, *entonces* $\mathbf{y} = c_1\mathbf{y}^1 + c_2\mathbf{y}^2$ *es una solución de*

$$\dot{\mathbf{x}}(t) = A(t)\,\mathbf{x}(t)+c_1\mathbf{f}^1(t)+c_2\mathbf{f}^2(t).$$

PRUEBA.
$$\begin{aligned}\dot{\mathbf{y}} &= c_1\dot{\mathbf{y}}^1+c_2\dot{\mathbf{y}}^2 = c_1A\mathbf{y}^1+c_1\mathbf{f}^1+c_2A\mathbf{y}^2+c_2\mathbf{f}^2\\ &= A[c_1\mathbf{y}^1+c_2\mathbf{y}^2]+c_1\mathbf{f}^1+c_2\mathbf{f}^2\\ &= A\mathbf{y}+c_1\mathbf{f}^1+c_2\mathbf{f}^2.\end{aligned}$$

En la hipótesis de que A y $\mathbf{f}$ sean continuas sobre $\langle -\infty, \infty\rangle$ (o sobre cualquier intervalo $\mathscr{I}$) se sabe que 9.1 tiene una solución única $\mathbf{x}$ sobre $\langle -\infty, \infty\rangle$ (o sobre $\mathscr{I}$) que satisface $\mathbf{x}(t_0)=\mathbf{x}_0(t_0\in\mathscr{I})$. Hemos mostrado esto para coeficientes constantes y aceptamos el resultado más general sin prueba. Supongamos que el sistema comienza en su posición de reposo $\mathbf{x}(0)=\mathbf{0}$. Entonces la ecuación 9.1 tiene una solución única $\mathbf{y}$ que satisface esta condición inicial. Considerando a $A(t)$ como una función matricial fija asociamos entonces con cada función $\mathbf{f}$ continua sobre $\langle -\infty, \infty\rangle$ la solución $\mathbf{y}$. Denotamos esta correspondencia por $L=\{(\mathbf{f},\mathbf{y})\}$ y

$$\mathbf{y} = L(\mathbf{f}).$$

L es una transformación definida sobre el espacio de funciones $\mathbf{f}$ continuas sobre $\langle -\infty, \infty\rangle$. La función $\mathbf{f}$ se llama a veces *entrada* (figura 5) y, la solución y, *salida*. El principio de superposición afirma que L es una transformación lineal:

$$L(c_1\mathbf{f}^1+c_2\mathbf{f}^2) = c_1L(\mathbf{f}^1)+c_2L(\mathbf{f}^2).$$

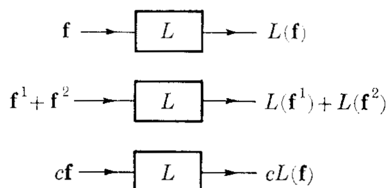


FIGURA 5

Es fácil de ver esto. Las funciones $\mathbf{y}^1=L(\mathbf{f}^1)$ y $\mathbf{y}^2=L(\mathbf{f}^2)$ son las soluciones de $\dot{\mathbf{x}}=A\mathbf{x}+\mathbf{f}^1$ y $\dot{\mathbf{x}}=A\mathbf{x}+\mathbf{f}^2$ que satisfacen $\mathbf{x}(0)=\mathbf{0}$. Por el principio de superposición $c_1\mathbf{y}^1+c_2\mathbf{y}^2$ es una solución de $\mathbf{x}=A\mathbf{x}^1+c_1\mathbf{f}^1+c_2\mathbf{f}^2$ y satisface $\mathbf{x}(0)=\mathbf{0}$. Por tanto, $L(c_1\mathbf{f}^1+c_2\mathbf{f}^2)=c_1\mathbf{y}^1+c_2\mathbf{y}^2=c_1L(\mathbf{f}^1)+c_2L(\mathbf{f}^2)$.

En términos de entradas y salidas el principio de superposición afirma que: *si las entradas son añadidas, las salidas se añaden* $(L(\mathbf{f}^1+\mathbf{f}^2)=L(\mathbf{f}^1)+L(\mathbf{f}^2))$; *si la entrada es multiplicada por una constante, la salida es multiplicada por la misma constante* $(L(c\mathbf{f})=cL(\mathbf{f}))$ (figura 5).

Si P es la solución matricial principal de

$$\dot{X}(t) = A(t)\,X(t),$$

sabemos (problema 4) que $\mathbf{y}=L(\mathbf{f})$ viene dada por

9.3
$$\mathbf{y}(t) = \int_0^t P(t)\,P^{-1}(s)\,\mathbf{f}(s)\,ds,$$

y en el caso particular en que A es una matriz constante

9.4 $$\mathbf{y}(t) = \int_0^t e^{A(t-s)}\mathbf{f}(s)\,ds.$$

Observemos una propiedad de la salida $\mathbf{y}$ que uno espera de un sistema físico real. El valor presente $\mathbf{y}(t)$ de la salida debe depender solamente de los valores pasados de la entrada. El sistema no puede prever la entrada. El sistema comienza en el tiempo $t=0$, el tiempo en que la entrada es aplicada, de modo que $\mathbf{f}(t)=\mathbf{0}$ para $t<0$. Luego $\mathbf{y}(t)=0$ para $t<0$ y en cualquier $t>0$ vemos, por 9.3, que *la salida depende solamente de los valores pasados de la entrada* $\mathbf{f}$.

Desde el punto de vista que estamos ahora tomando podemos ver una diferencia entre los coeficientes constantes y los coeficientes no constantes. El sistema lineal está caracterizado por la matriz $A(t)$. Si A es constante tenemos un "sistema estacionario"; no cambia con el tiempo. Con coeficientes no constantes el sistema cambia. Así pues, para un sistema estacionario, esperamos, independientemente de cuando sea el momento en que la entrada se aplique, puesto que la salida debe ser la misma. Se dice a veces que "el sistema es invariante respecto a un cambio en el origen del tiempo". En forma más precisa lo que se quiere decir es que si $\mathbf{y}=L(\mathbf{f})$ y $\mathbf{y}^*=L(\mathbf{f}^*)$, donde $\mathbf{f}^*(t)=\mathbf{f}(t-\delta)$ y $\mathbf{f}(t)=\mathbf{0}$ para $t<0$, entonces $\mathbf{y}^*(t)=\mathbf{y}(t-\delta)$. *Si la entrada de un sistema estacionario se cambia en el tiempo en la cantidad $\delta>0$, entonces la salida es cambiada en el tiempo la misma cantidad.* Esto sigue fácilmente de 9.4 por un simple cambio de variable en la integral:

$$\mathbf{y}^*(t) = \int_0^t e^{A(t-s)}\mathbf{f}(s-\delta)\,ds = \int_\delta^t e^{A(t-s)}\mathbf{f}(s-\delta)\,ds$$
$$= \int_0^{t-\delta} e^{A(t-\delta-u)}\mathbf{f}(u)\,du = \mathbf{y}(t-\delta).$$

En términos de la solución matricial principal P la diferencia entre coeficientes constantes y coeficientes no constantes es ésta: para A una matriz constante

$$P(t)\,P^{-1}(s) = P(t-s);$$

para matrices no constantes esto no es cierto (véase el problema 5).

Nota. En la anterior discusión nos limitamos a funciones $\mathbf{f}$ continuas sobre $\langle-\infty,\infty\rangle$. Todo lo que aquí hemos dicho puede extenderse a funciones integrables sobre $\langle-\infty,\infty\rangle$. Si $\mathbf{f}$ es integrable sobre $\langle-\infty,\infty\rangle$, entonces $\mathbf{x}$ se dice que es una *solución* de $\dot{\mathbf{x}}=A\mathbf{x}+\mathbf{f}$ sobre $\langle-\infty,\infty\rangle$ si $\mathbf{x}$ es continua sobre $\langle-\infty,\infty\rangle$ y satisface la ecuación diferencial en todos los puntos en que $\mathbf{f}$ es continua.

Si g es integrable sobre $\langle -\infty, \infty\rangle$ la siguiente extensión del teorema fundamental del cálculo es cierta: si $G(t) = \int_0^t g(s)\,ds$, entonces G es continua sobre $\langle -\infty, \infty\rangle$ y en cada t donde g es continua $G'(t) = g(t)$.

Se sigue de ello que

$$\mathbf{x}(t) = P(t)\mathbf{c} + P(t)\int_0^t P^{-1}(s)\mathbf{f}(s)\,ds$$

es la solución sobre $\langle -\infty, \infty\rangle$ de $\dot{\mathbf{x}}(t) = A(t)\,\mathbf{x}(t) + \mathbf{f}(t)$ que satisface $\mathbf{x}(0) = \mathbf{c}$. En particular, no debe tenerse duda alguna en cuanto a lo correcto de aplicar lo que aquí hemos aprendido a entradas continuas a trozos (problema 3, pág. 635).

Debe tenerse presente que la ecuación diferencial lineal de segundo orden

9.5 $$\ddot{x}(t) + b(t)\,\dot{x}(t) + c(t)\,x(t) = f(t)$$

es un caso especial del sistema 9.1. Un sistema equivalente es

9.6 $$\begin{aligned}\dot{x}(t) &= y(t)\\ \dot{y}(t) &= -c(t)\,x(t) - b(t)\,y(t) + f(t).\end{aligned}$$

Aquí $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}$, $A = \begin{pmatrix} 0 & 1 \\ -c & -b \end{pmatrix}$, y $\mathbf{f} = \begin{pmatrix} 0 \\ f \end{pmatrix}$. La solución x de 9.5 es el primer componente de $\mathbf{x}$. Así pues, el teorema de superposición nos dice que, si x_i es una solución de $\ddot{x}(t) + b(t)\,\dot{x}(t) + c(t)\,x(t) = f_i(t)$, entonces $\sum_{i=1}^{n} c_i x_i$ es una solución de $\ddot{x}(t) + b(t)\,\dot{x}(t) + c(t)\,\dot{x}(t) = \sum_{i=1}^{n} c_i f_i(t)$.

Para coeficientes constantes b y c, correspondiéndose con 9.4, tenemos que

9.7 $$x(t) = \int_0^t w(t-s)\,f(s)\,ds$$

es la salida para una entrada f, donde w es la solución de $\ddot{x}(t) + b\dot{x}(t) + cx(t) = 0$ que satisface $w(0) = 0$, $\dot{w}(0) = 1$.

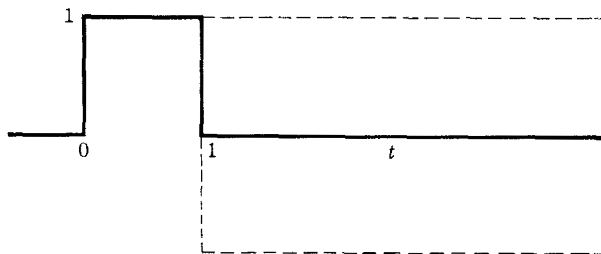


FIGURA 6

9.8 Ejemplo. Determínese la respuesta (la salida) de

$$\ddot{x}+6\dot{x}+5x = f$$

al impulso de entrada cuadrado

$$f(t) = \begin{cases} 0, & t < 0 \\ 1, & 0 \leqslant t < 1 \\ 0, & t \geqslant 1 . \end{cases}$$

SOLUCIÓN. La ecuación característica es $\lambda^2+6\lambda+5=0$, y las raíces características son $\lambda = -1, -5$. Encontramos entonces que

$$w(t) = \tfrac{1}{4}(e^{-t}-e^{-5t}).$$

Sea u la función de escalón unitario ($u(t)=0$, $t<0$, $u(t)=1, t \geqslant 0$). Entonces el pulso cuadrado $f(t)=u(t)-u(t-1)$ (figura 6). La respuesta y a la función de escalón unitario es (problema 7, pág. 636): $y(t)=0$ para $t<0$ y

$$\begin{aligned} y(t) &= \int_0^t w(t-s)\,u(s)\,ds = \int_0^t w(s)\,u(t-s)\,ds = \int_0^t w(s)\,ds \\ &= \tfrac{1}{4}\int_0^t (e^{-s}-e^{-5s})\,ds \\ &= \tfrac{1}{5}-\tfrac{1}{4}e^{-t}+\tfrac{1}{20}e^{-5t} \quad \text{para} \quad t \geqslant 0 . \end{aligned}$$

Este es un sistema estacionario (de coeficientes constantes). Luego la respuesta y^* correspondiente a la entrada $u(t-1)$ es $y^*(t)=y(t-1)$. Por tanto, la respuesta x al impulso cuadrado es

$$x(t) = y(t)-y(t-1).$$

Esta salida se ha representado en la figura 7.

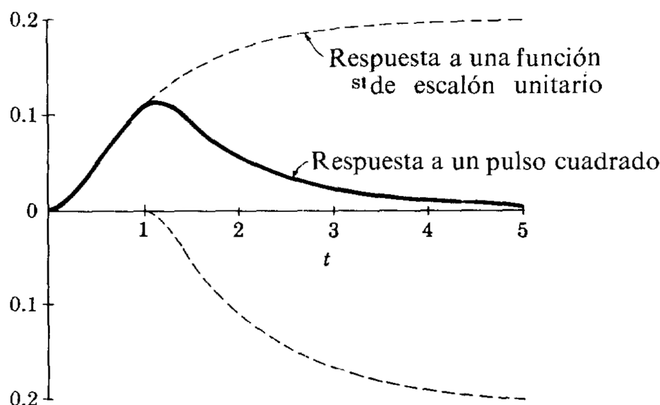


FIGURA 7

Algunas consecuencias simples del principio de superposición son:

9.9 Corolario. *Cualquier combinación lineal de soluciones de la ecuación homogénea* $\mathbf{x}(t) = A(t)\,\mathbf{x}(t)$ *es también una solución.*

PRUEBA. Hágase $\mathbf{f}^1 = \mathbf{f}^2 = \mathbf{0}$ en el teorema 9.2.

9.10 Corolario. *La solución general de la ecuación completa* $\dot{\mathbf{x}}(t) = A(t)\,\mathbf{x}(t)+\mathbf{f}(t)$ *es la solución general de la ecuación homogénea más cualquier solución particular de la ecuación completa.*

PRUEBA. Sea $\mathbf{y}$ una solución (particular) de $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$ y sea $\mathbf{x}$ cualquier otra solución. Entonces, según el principio de superposición $\mathbf{x}-\mathbf{y}$ es una solución de la ecuación homogénea. De donde $\mathbf{x}(t)-\mathbf{y}(t) = P(t)\mathbf{c}$ y $\mathbf{x}(t) = P(t)\mathbf{c}+\mathbf{y}(t)$. De nuevo, según el principio de superposición, toda función de esta forma es una solución de la ecuación completa, de donde ésta es la solución general.

Problemas

1. Determínense las respuestas x y $\dot{x}$ de

a) $\ddot{x}+10.1\,\dot{x}+x = f$
b) $\ddot{x}+110\dot{x}+1\,000x = f$

donde f es el pulso cuadrado del ejemplo 9.8. Grafíquense las respuestas x y $\dot{x}$.

2. Resuélvase el ejemplo 9.8 con

a) $f(t) = \begin{cases} 0, & t < 0 \\ 1, & 0 \leqslant t < \frac{1}{5} \\ 0, & \frac{1}{5} \leqslant t \end{cases}$ *b*) $f(t) = \begin{cases} 0, & t < 0 \\ 1, & 0 \leqslant t < 5 \\ 0, & 5 \leqslant t\,. \end{cases}$

3. Determínese la respuesta x de $\ddot{x}+10\dot{x}+24x = f$ donde

a) $f(t) = \begin{cases} 0, & t \leqslant 0 \\ t, & 0 \leqslant t \end{cases}$ *b*) $f(t) = \begin{cases} 0, & t \leqslant 0 \\ t, & 0 \leqslant t \leqslant 1 \\ 1, & 1 \leqslant t \end{cases}$

c) $f(t) = \begin{cases} 0, & t \leqslant 0 \\ t, & 0 \leqslant t \leqslant 1 \\ 2-t, & 1 \leqslant t \leqslant 2 \\ 0, & 2 \leqslant t\,. \end{cases}$

Dibújense las gráficas de las entradas y salidas.

4. Suponiendo la unicidad de las soluciones, pruébese que si $P(t)$ es la solución matricial principal de $\dot{X}(t) = A(t)\,X(t)$, entonces 9.3 es la solución de $\dot{\mathbf{x}}(t) = A(t)\,\mathbf{x}(t)+\mathbf{f}(t)$ que satisface $\mathbf{x}(0) = \mathbf{0}$.

5. Supongamos que $A(t)$ es continua para todo t. Sea P la solución matricial principal de $\dot{X} = A(t)X$. Pruébese que, si $P(t)P^{-1}(s) = P(t-s)$ para toda t y toda s, entonces $A(t)$ es una matriz constante.

6. Si $i\omega$ no es un valor característico de A, pruébese que

$$\mathbf{x}(t) = e^{At}\mathbf{c}+(i\omega I-A)^{-1}\,\mathbf{b}e^{i\omega t}$$

es la solución general de

$$\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{b}e^{i\omega t}\,.$$

Conclúyase de aquí que, si A no tiene valores característicos imaginarios puros, entonces

$$\mathbf{x}(t) = e^{At}\mathbf{c} + \sum_{k=1}^{n} (i\omega_k I-A)^{-1}\,\mathbf{b}^k e^{i\omega_k t}$$

es la solución general de

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + \sum_{k=1}^{n} \mathbf{b}^k e^{i\omega_k t}\,.$$

7. Determínese una solución particular de cada una de las siguientes ecuaciones diferenciales:

a) $\dot{x}(t) = 2x(t)-y(t)$, $\dot{y}(t) = x(t)+y(t)+4\cos 3t$

b) $\dot{x} = x+y-2\cos$, $\dot{y} = x-y-4\operatorname{sen}$

c) $\dot{x}(t) = x(t)+y(t)-\cos 2t$, $\dot{y}(t) = x(t)-y(t)+\operatorname{sen} 3t$

d) $\ddot{x}(t)+5\dot{x}(t)+3x(t) = \cos 10t$

e) $3\ddot{x}(t)+7\dot{x}(t)+9x(t) = \cos t-\frac{1}{3}\cos 3t$.

10. OSCILACIONES LINEALES. $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$

Al estudiar el sistema lineal

10.1 $$\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$$

A ha sido una matriz compleja y $\mathbf{f}$ una función vectorial compleja. La principal razón para tomar como sistema numérico el campo complejo fue la de introducir valores característicos y vectores característicos. Nuestro interés, sin embargo, recae primariamente en los sistemas reales y en las soluciones reales. Supondremos por ello en toda esta sección que A es una matriz constante real y $\mathbf{f}$ una función vectorial real continua sobre $\langle -\infty, \infty\rangle$. Lo que queremos es investigar en forma más detallada que anteriormente

el carácter de las soluciones reales de 10.1. Nos interesaremos particularmente en las soluciones periódicas.

Comenzaremos por discutir las soluciones reales de la ecuación homogénea

10.2 $$\dot{\mathbf{x}} = A\mathbf{x}.$$

Sean λ_1 y λ_2 los valores característicos de A. Pueden, desde luego, ser reales o complejos. Si son complejos, entonces, como A es real, son complejos conjugados uno de otro. El carácter de las soluciones depende de la naturaleza de λ_1 y λ_2.

Caso 1. λ_1, λ_2 reales y distintos ($\lambda_1 \neq \lambda_2$).

Designemos por λ_1 a la mayor de las dos raíces ($\lambda_1 > \lambda_2$), y sean $\mathbf{v}^1$ y $\mathbf{v}^2$ vectores característicos correspondientes. Como A y sus valores característicos son reales, los vectores característicos son reales. Supondremos, lo que desde luego podemos hacer, que los vectores característicos $\mathbf{v}^1$ y $\mathbf{v}^2$ son unitarios. La solución general de 10.2 es

10.3 $$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{v}^1 + c_2 e^{\lambda_2 t}\mathbf{v}^2.$$

Toda solución $\mathbf{x}$ de 10.2 define una curva γ en el plano. Queremos investigar primero lo que sucede a la dirección de γ cuando $t \to \infty$ y cuando $t \to -\infty$. Si $A\mathbf{x}^0 = 0$, entonces el punto $\mathbf{x}^0$ se llama *estado de equilibrio* o *punto crítico* del sistema 10.2. Así pues, si $\mathbf{x}^0$ es un estado de equilibrio $\mathbf{x} = \mathbf{x}^0$ es una solución de 10.2 y la curva solución γ es simplemente un punto. Si λ_1 y λ_2 son distintas de cero, entonces A es no singular y el único estado de equilibrio es el origen. Si $\lambda_1 = 0$, entonces $A\mathbf{v}^1 = \mathbf{0}$ y todo punto sobre la recta que pasa por el origen y tiene la dirección de $\mathbf{v}^1$ es un estado de equilibrio. En cualquier punto que no sea un estado de equilibrio, la curva solución γ tiene un vector tangente unitario

$$\frac{\dot{\mathbf{x}}(t)}{|\dot{\mathbf{x}}(t)|} = \frac{c_1\lambda_1 e^{\lambda_1 t}\mathbf{v}^1 + c_2\lambda_2 e^{\lambda_2 t}\mathbf{v}^2}{(c_1{}^2\lambda_1{}^2 e^{2\lambda_1 t} + 2c_1c_2\lambda_1\lambda_2 e^{(\lambda_1+\lambda_2)t}\mathbf{v}^1\cdot\mathbf{v}^2 + c_2{}^2\lambda_2{}^2 e^{2\lambda_2 t})^{1/2}}$$

$$= \frac{c_1\lambda_1\mathbf{v}^1 + c_2\lambda_2 e^{(\lambda_2-\lambda_1)t}\mathbf{v}^2}{(c_1{}^2\lambda_1{}^2 + 2c_1c_2\lambda_1\lambda_2 e^{(\lambda_2-\lambda_1)t}\mathbf{v}^1\cdot\mathbf{v}^2 + c_2{}^2\lambda_2{}^2 e^{2(\lambda_2-\lambda_1)t})^{1/2}}.$$

Como $\lambda_1 > \lambda_2$, vemos que

i) *Si* $c_1\lambda_1 \neq 0$, *la tangente unitaria a la curva solución* γ *tiende a* $\pm\mathbf{v}^1$ *cuando* $t \to \infty$.

Análogamente,

ii) *Si* $c_2\lambda_2 \neq 0$, *la tangente unitaria a la curva solución* γ *tiende a* $\pm\mathbf{v}^2$ *cuando* $t \to -\infty$.

Esto está ilustrado en las figuras 8 y 9. El significado de $c_1 \lambda_1 = 0$ es claro. Si $c_1 = 0$ la solución comienza sobre la recta $\mathscr{L}_2$ a través del origen en la dirección $\mathbf{v}^2$ (ecuación 10.3) y permanece en esta recta. Si $\lambda_1 = 0$, entonces todas las curvas solución son rectas paralelas a $\mathbf{v}^2$.

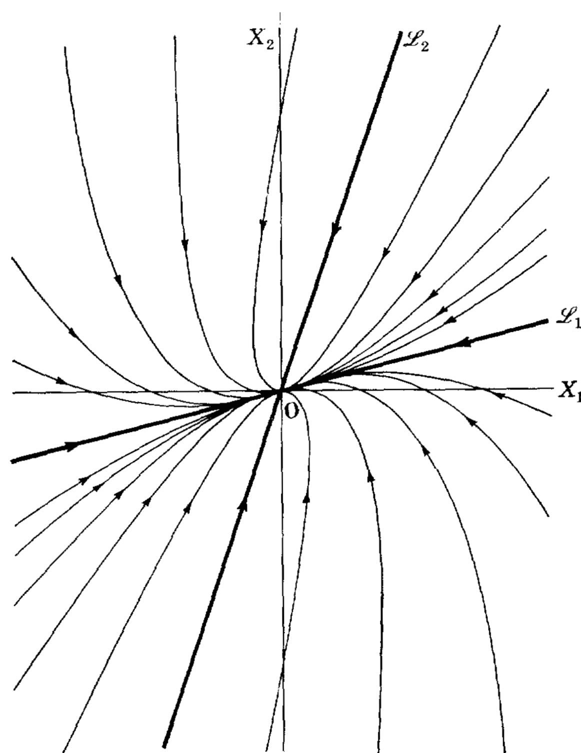


FIGURA 8 Nudo estable ($\lambda_2 < \lambda_1 < 0$)

1 *a*) *Valores característicos negativos* ($\lambda_2 < \lambda_1 < 0$). Se sigue entonces de 10.3 que

iii) *Toda solución* $\mathbf{x}(t) \to 0$ *cuando* $t \to \infty$.

El estado de equilibrio en el origen se dice entonces que es "asintóticamente estable". No importa hasta qué punto el sistema esté perturbado, siempre tiende a volver al estado de equilibrio. Además

iv) $|\mathbf{x}(t)| \to \infty$ *cuando* $t \to -\infty$ ($\mathbf{x}(0) \neq \mathbf{0}$).

Con la información obtenida (i-iv) podemos dibujar las curvas solución (figura 8). En este caso, el estado de equilibrio $\mathbf{0}$ se llama *nudo* o *nodo estable*.

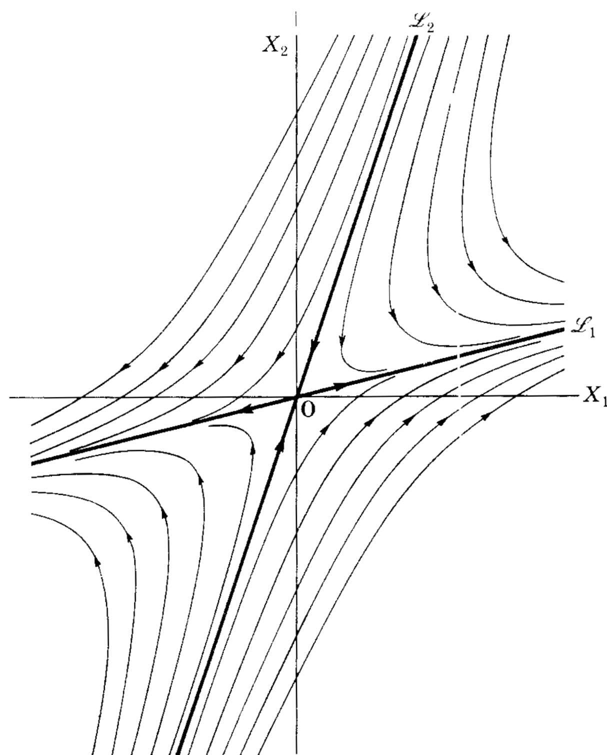


FIGURA 9 Punto de ensilladura ($\lambda_2 < 0 < \lambda_1$)

1 *b*) *Valores característicos positivos* ($\lambda_1 > \lambda_2 > 0$). Es análogo al 1 *a*) si reemplazamos t por $-t$. Las flechas en la figura 8 tienen que invertir su dirección. Todo lo cerca del origen que queramos hay soluciones que dejan la vecindad del origen (cualquier vecindad dada). El estado de equilibrio se dice que es "inestable" y se llama *nudo o nodo inestable.*

1 *c*) *Un valor característico positivo y otro negativo* ($\lambda_2 < 0 < \lambda_1$). Vemos, por 10.3, que hay solamente dos curvas solución que tienden al origen. Éstas se corresponden con $c_1 = 0$ y $c_2 > 0$ o $c_2 < 0$. Todas las otras soluciones tienden a infinito cuando $t \to \infty$. El origen es "inestable" y se llama *punto de ensilladura* (figura 9).

Caso 2. Raíces complejas. $\lambda_1 = \beta + i\omega$, $\lambda_2 = \beta - i\omega$.

Aquí en el plano, que es real, no tenemos ya vectores reales característicos. Con A real las raíces son complejas conjugadas una de

la otra y así lo son los vectores característicos. De donde las soluciones reales están dadas por

$$\mathbf{x}(t) = c_1\mathbf{v}^1 e^{(\beta+i\omega)t} + \bar{c}_1\bar{\mathbf{v}}^1 e^{(\beta-i\omega)t}$$

$$= 2\,\mathrm{Rl}(c_1\mathbf{v}^1 e^{(\beta+i\omega)t})$$

donde c_1 es un número complejo arbitrario. Sea $\mathbf{v}^1 = \mathbf{u}+i\mathbf{v}$ donde $\mathbf{u}$ y $\mathbf{v}$ son vectores reales linealmente independientes (problema 3), y $2c_1 = ae^{i\delta}$. Entonces la solución general real es

10.4 $\qquad \mathbf{x}(t) = ae^{\beta t}[\cos(\omega t+\delta)\mathbf{u} - \operatorname{sen}(\omega t+\delta)\mathbf{v}].$

De donde, cuando $\omega t+\delta = k\pi$, la solución está en la dirección $\mathbf{u}$. Cuando $\omega t+\delta = \frac{1}{2}(2k+1)\pi$ la solución está en la dirección $\mathbf{v}$. Los componentes en las direcciones $\mathbf{u}$ y $\mathbf{v}$ oscilan y están 90° fuera de fase.

2*a*) *Raíces imaginarias puras* ($\lambda_1 = i\omega$, $\lambda_2 = -i\omega$). Aquí

$$\mathbf{x}(t) = a[\cos(\omega t+\delta)\mathbf{u} - \operatorname{sen}(\omega t+\delta)\mathbf{v}],$$

y toda solución es periódica con periodo $2\pi/\omega$. Las curvas solución son curvas cerradas. Si $\mathbf{u}$ y $\mathbf{v}$ fueran ortogonales, serían elipses con centro en el origen. De otra forma, son elipses distorsionadas (figura 10). Las soluciones que comienzan cerca del origen permanecen cerca de él. El origen es "estable" y se llama *centro*.

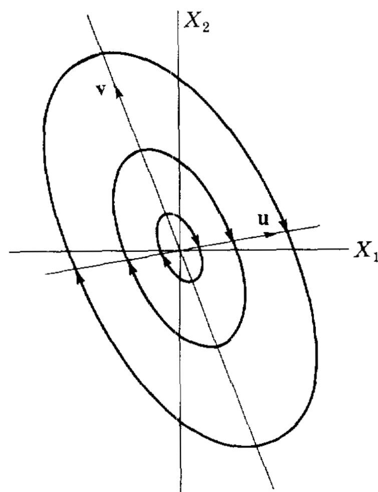


FIGURA 10 Centro ($\lambda_1 = i\omega = -\lambda_2$)

2*b*) *Partes reales negativas* ($\beta < 0$). Observando 10.4 vemos que $\mathbf{x}(t) \to \mathbf{0}$ cuando $t \to \infty$. Las intersecciones sucesivas con las rectas que parten del origen en las direcciones $\mathbf{u}$ y $\mathbf{v}$ se mueven hacia el origen a medida que el tiempo aumenta, y las curvas solución se enroscan en

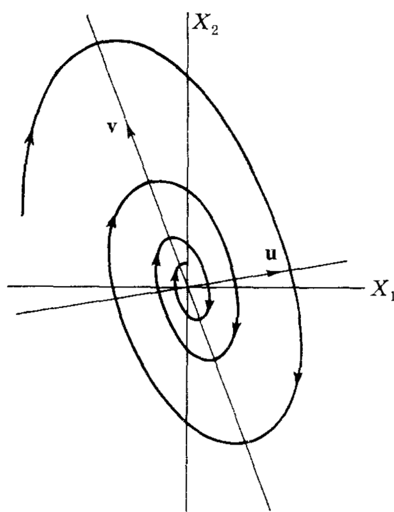


FIGURA 11 Foco estable ($\beta < 0$)

espiral alrededor del origen. El origen es asintóticamente estable y se llama *foco estable* (figura 11).

2*c*) *Partes reales positivas* ($\beta < 0$). Aquí las soluciones se enroscan en espiral alrededor y hacia afuera del origen. El origen es inestable y se llama *foco inestable*. Las curvas solución son como las de la figura 11 con las flechas en dirección contraria.

Caso 3. Raíces iguales ($\lambda_1 = \lambda_2$).

Como $\lambda_1 = \lambda_2$, las raíces son reales. Si las raíces son negativas, todas las soluciones tienden a cero cuando $t \to \infty$. El origen es asintóticamente estable y el cuadro no es muy diferente del que aparece en la figura 8 (figura 12). Las soluciones no oscilan y el origen es una vez más un *módulo estable*. Si las raíces son positivas, el origen es inestable y se llama *nodo inestable*.

Oscilaciones forzadas. Deseamos discutir las soluciones periódicas de la ecuación completa 10.1 cuando $\mathbf{f} \neq \mathbf{0}$. Las soluciones periódicas se llaman "oscilaciones forzadas". Sea $\mathbf{g}$ una función vectorial definida sobre $\langle -\infty, \infty\rangle$. La función $\mathbf{g}$ se dice que es periódica si para algún $T \neq 0$, $\mathbf{g}(t+T) = \mathbf{g}(t)$ para todo t en $\langle -\infty, \infty\rangle$. El número T se llama periodo de $\mathbf{g}$. Las funciones periódicas *triviales* son las funciones constantes. Si $\mathbf{g}$ es una solución periódica no trivial, tiene un periodo positivo mínimo T. Llamaremos a tal periodo positivo mínimo, simplemente *periodo mínimo*.

Hemos visto que la ecuación homogénea $\dot{\mathbf{x}} = A\mathbf{x}$ tiene soluciones periódicas no triviales si y sólo si las raíces características de A son imaginarias puras. Cuando $\lambda = i\omega$, todas las soluciones (salvo la solución trivial)

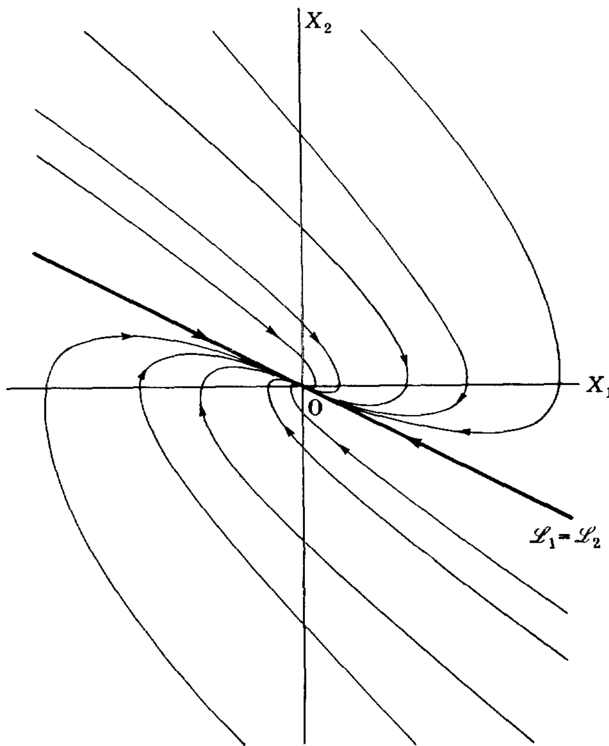


FIGURA 12 Nodo estable ($\lambda_1 = \lambda_2 < 0$)

son periódicas con periodo mínimo $2\pi/\omega$. Éstas son las llamadas "oscilaciones libres" del sistema. Queremos ahora discutir las oscilaciones forzadas. Son éstas, como hemos dicho, las soluciones periódicas de

10.5 $$\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}, \quad \mathbf{f} \neq \mathbf{0}.$$

Observamos primero que si debemos tener soluciones periódicas necesitamos suponer que el término forzante **f** es periódico.

10.6 Lema. *Si $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$ tiene una solución periódica de periodo mínimo T_0, entonces* **f** *es periódica de periodo T_0. Si T es el periodo mínimo de* **f**, *entonces $T_0 = kT$ para algún entero positivo k.*

PRUEBA. Sea **x** una solución periódica de 10.5 de periodo mínimo T_0. Entonces $\dot{\mathbf{x}}$ es también periódica de periodo T_0. Como $\mathbf{f}(t) = \dot{\mathbf{x}}(t) - A\mathbf{x}(t)$, se sigue que **f** es periódica de periodo T_0. Si **f** es periódica de periodo mínimo T, entonces sus únicos periodos positivos son T, $2T$, ..., kT, ... Por tanto, para algún entero positivo k, $T_0 = kT$. Y esto completa la prueba.

En conexión son este lema, si $T_0 = kT$ con $k > 1$ la solución periódica se llama *subarmónica de orden k*. Se dice que es "sub" porque la frecuencia de la solución periódica es $\frac{1}{T_0} = \frac{1}{kT}$, que es un k-ésimo de la frecuencia del término forzante. Por ejemplo, la solución general de

10.7 $$\ddot{x} + x = \cos 2t$$

es $a \cos(t+\delta) - \frac{1}{3}\cos 2t$. El periodo mínimo del término forzante $\cos 2t$ es π. Si $a \neq 0$, la solución tiene un periodo mínimo igual a 2π, y estas soluciones son subarmónica de orden 2. La única solución de periodo π es $-\frac{1}{3}\cos 2t$, que corresponde a $a = 0$.

Si 10.5 ha de tener soluciones periódicas, de acuerdo con el lema 10.6 debemos suponer que **f** es periódica. Pero esto en sí no implica, sin embargo, que habrá soluciones periódicas. Un ejemplo es

$$\ddot{x} + x = \cos t.$$

La solución general es $a \cos(t+\delta) + \frac{1}{2}t \operatorname{sen} t$, y no hay soluciones periódicas. Lo que es particular en este ejemplo es que hay una oscilación libre cuyo periodo es el mismo que el del término forzante. Lo que queremos demostrar es que si no es éste el caso, siempre hay soluciones periódicas. Llegamos a este resultado a través de varios lemas que en sí mismos son de interés.

10.8 Lema. *La ecuación homogénea $\dot{\mathbf{x}} = A\mathbf{x}$ tiene una solución periódica distinta de cero de periodo T si y sólo si $I - e^{AT}$ es singular.*

PRUEBA. La solución general es $\mathbf{x}(t) = e^{At}\mathbf{c}$. Hay entonces una solución periódica de periodo T si y sólo si $e^{A(t+T)}\mathbf{c} = e^{AT}\mathbf{c}$ para todo t o $e^{At}e^{AT}\mathbf{c} = e^{At}I\mathbf{c}$ para toda t. Por tanto, hay una solución periódica distinta de cero si y sólo si la ecuación

$$(I - e^{AT})\mathbf{c} = \mathbf{0}$$

tiene una solución distinta de cero $\mathbf{c}$. Esto es cierto si y sólo si $I - e^{AT}$ es singular.

Lo que nos interesa respecto a lo que acabamos de afirmar no es que nos dé un criterio sobre la existencia de soluciones periódicas de $\dot{\mathbf{x}} = A\mathbf{x}$. Esto ya era conocido. Lo que es importante es que, como sabemos acerca de soluciones periódicas de $\dot{\mathbf{x}} = A\mathbf{x}$, el lema 10.8 nos da un criterio sobre la no singularidad de $I - e^{At}$.

10.9 Lema. *Supongamos que* **f** *es periódica de periodo T. Entonces una solución* $\mathbf{x}$ *de* $\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{f}$ *es periódica de periodo T si y sólo si* $\mathbf{x}(0) = \mathbf{x}(T)$.

PRUEBA. La necesidad de la condición es trivial; si $\mathbf{x}$ es periódica de periodo T, entonces $\mathbf{x}(t+T) = \mathbf{x}(t)$ para todo t, y por tanto, $\mathbf{x}(T) = \mathbf{x}(0)$.

Para probar la suficiencia de la condición supongamos que $\mathbf{x}$ es una solución y que $\mathbf{x}(0) = \mathbf{x}(T)$. Definamos $\mathbf{y}(t) = \mathbf{x}(t+T)$. Entonces

$$\begin{aligned}\dot{\mathbf{y}}(t) &= \frac{d}{dt}\mathbf{x}(t+T) = \dot{\mathbf{x}}(t+T)\\ &= A\mathbf{x}(t+T)+\mathbf{f}(t+T)\\ &= A\mathbf{y}(t)+\mathbf{f}(t),\end{aligned}$$

y $\mathbf{y}$ es también una solución. (Aquí hemos usado el hecho de que $\mathbf{f}$ es periódica de periodo T.) Como $\mathbf{y}(0) = \mathbf{x}(T) = \mathbf{x}(0)$, tenemos, según la unicidad de las soluciones, que $\mathbf{x}(t+T) = \mathbf{y}(t) = \mathbf{x}(t)$ para toda t. Y esto completa la prueba.

Llegamos ahora al resultado principal.

10.10 Teorema. *Sea* $\mathbf{f}$ *continua y periódica de periodo* T. *La ecuación* $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$ *tiene una solución periódica única de periodo* T *si y sólo si la ecuación* $\dot{\mathbf{x}} = A\mathbf{x}$ *no tiene solución periódica alguna distinta de cero de periodo* T.

PRUEBA. La solución general de $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$ es

$$\mathbf{x}(t) = e^{At}\mathbf{c} + \int_0^t e^{A(t-s)}\mathbf{f}(s)\,ds.$$

De acuerdo con el lema 10.9, hay una solución periódica de periodo T si y sólo si

$$\mathbf{c} = e^{AT}\mathbf{c} + \int_0^T e^{A(T-s)}\mathbf{f}(s)\,ds;$$

es decir, si y sólo si para algún $\mathbf{c}$

$$(I-e^{AT})\mathbf{c} = \int_0^T e^{A(T-s)}\mathbf{f}(s)\,ds.$$

Así pues, hay una solución periódica única si y sólo si la anterior ecuación tiene una solución única $\mathbf{c}$. Esto es cierto si y sólo si $I-e^{AT}$ es no singular, y la conclusión de este teorema se sigue del lema 10.8.

Nótese que el anterior teorema implica la existencia y unicidad de una solución periódica *de periodo T*. Por ejemplo, la ecuación diferencial homogénea correspondiente a 10.7 no tiene ninguna solución periódica distinta de cero de periodo π. La ecuación completa tiene una solución periódica única de periodo π, pero una infinidad de soluciones periódicas. Tenemos, sin embargo, como una sencilla consecuencia del anterior teorema y del lema 10.6:

10.11 Corolario. *Sea* $\mathbf{f}$ *continua y periódica con periodo mínimo* T. *Si* $\dot{\mathbf{x}} = A\mathbf{x}$ *no tiene ninguna solución periódica de periodo* kT *para ningún entero positivo* k, *entonces* $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$ *tiene una solución periódica única, y esta solución periódica tiene periodo* T.

PRUEBA. Según el lema 10.6 las únicas soluciones periódicas posibles son las de periodo kT donde k es un entero positivo. De acuerdo con el teorema 10.10 hay, para cada $k = 1, 2, \ldots$ una solución periódica única de periodo kT. De donde, como la solución periódica de periodo T tiene también periodo kT para cualquier entero positivo k, ésta es la única solución periódica posible.

Las condiciones de este corolario son ciertamente satisfechas si las raíces características de A tienen partes reales no nulas. También se satisfacen si hay raíces características imaginarias puras $\pm i\omega(\omega > 0)$, pero $2\pi/\omega \neq kT$ para $k = 1, 2, \ldots$ Sin embargo, en este último caso de raíces imaginarias puras hay una infinidad de lo que se llaman soluciones "casi periódicas" que se comportan en cierta medida como soluciones periódicas.

Supongamos que las raíces características de A tienen partes reales distintas de cero. Entonces hay una solución periódica única de $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$ y esta solución periódica tiene periodo T. (Aun estamos suponiendo, desde luego, que $\mathbf{f}$ es periódica de periodo T.) Denotemos esta solución periódica por $\bar{\mathbf{x}}(t)$. Si hay una raíz característica de A con parte real positiva, entonces hay soluciones $\mathbf{y}$ de $\dot{\mathbf{x}} = A\mathbf{x}$ con la propiedad de que $|\mathbf{y}(t)| \to \infty$ cuando $t \to \infty$. Ahora bien, $c\mathbf{y}+\bar{\mathbf{x}}$ es una solución de $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$ para toda $\mathbf{c}$. Así pues, arbitrariamente próximas a la solución periódica hay soluciones que no permanecen próximas, y la solución periódica se dice que es "inestable". Si todas las raíces características tienen partes reales negativas, entonces toda solución $\mathbf{y}$ de $\dot{\mathbf{x}} = A\mathbf{x}$ tiene la propiedad de que $\mathbf{y}(t) \to \mathbf{0}$ cuando $t \to \infty$. Toda solución de $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$ es de la forma $\mathbf{y}+\mathbf{x}$, de donde toda solución tiende a la solución periódica $\bar{\mathbf{x}}$ cuando $t \to \infty$. Esta solución periódica $\bar{\mathbf{x}}$ se llama, entonces, *solución de estado estable*. Así pues, si $\mathbf{x}$ es una solución cualquiera que satisface $\mathbf{x}(0) = \mathbf{x}^0$, entonces $\mathbf{x} = \mathbf{y}+\bar{\mathbf{x}}$, donde $\mathbf{y}$ es la solución de $\dot{\mathbf{x}} = A\mathbf{x}$ que satisface $\mathbf{y}(0) = \mathbf{x}^0 - \bar{\mathbf{x}}(0)$. A la solución $\mathbf{y}$ se le llama *transitoria*, ya que tiende a cero cuando $t \to \infty$ y describe la forma en que la solución se aproxima a la solución de estado estable.

10.12 Ejemplo. Determínese la solución de estado estable de

$$\ddot{x}+2\alpha\dot{x}+x = \cos t,$$

cuando *a*) $\alpha = 100$, *b*) $\alpha = 0.001$.

SOLUCIÓN. Para un término forzante sinusoidal es particularmente fácil

encontrar la solución de estado estable. Consideremos la ecuación más general

10.13 $$\ddot{x}+2\alpha\dot{x}+\omega_0^2 x = \cos\omega t.$$

El polinomio característico es $\varphi(\lambda) = \lambda^2+2\alpha\lambda+\omega_0^2$, y vemos para $\alpha > 0$ y $\omega_0^2 > 0$ que las raíces características tienen partes reales negativas y hay una solución de estado estable. Una sencilla forma de determinar la solución de estado estable es considerar la ecuación

10.14 $$\ddot{x}+2\alpha\dot{x}+\omega_0^2 x = e^{i\omega t},$$

y buscar una solución particular de la forma $ae^{i\omega t}$. Sustituyendo en la ecuación nos da

$$ae^{i\omega t}((i\omega)^2+2\alpha(i\omega)+\omega_0^2) = a\varphi(i\omega)e^{i\omega t} = e^{i\omega t}.$$

Como no hay raíces características imaginarias, $\varphi(i\omega) \neq 0$ y

$$ae^{i\omega t} = \frac{1}{\varphi(i\omega)}\, e^{i\omega t}.$$

Ésta es entonces la solución de estado estable de 10.14 (podemos ahora decir que esto es una consecuencia del teorema de superposición) y la solución de estado estable $x(t)$ de 10.13 es

$$x(t) = \mathrm{Rl}\left(\frac{1}{\varphi(i\omega)}\, e^{i\omega t}\right).$$

Para la ecuación particular de este ejemplo ($\omega = \omega_0 = 1$)

$$x(t) = \mathrm{Rl}\left(\frac{1}{2\alpha i}\, e^{it}\right) = \frac{1}{2\alpha}\operatorname{sen} t.$$

La salida tiene la misma frecuencia que la entrada (a esto se le llama *resonancia*), y la salida está 90° fuera de fase respecto a la entrada. La derivada x está en fase con la entrada. La amplitud de la solución de estado estable depende inversamente de α. Las soluciones son:

a) $x(t) = 0.005 \operatorname{sen} t$. *b*) $x(t) = 500 \operatorname{sen} t$.

Problemas

1. Determínese para cada uno de los siguientes sistemas la naturaleza del equilibrio estable (nodo estable, foco estable, punto de ensilladura, etc.)

a) $\dot{x} = 2x+4y$, $\dot{y} = 5x+3y$

b) $\dot{x} = x-y$, $\dot{y} = 3x+7y$

c) $\dot{x} = 3x-2y$, $\dot{y} = -4x+7y$

d) $\dot{x} = x-5y$, $\dot{y} = 12x-y$

e) $\dot{x} = 2x-3y$
$\dot{y} = 6x-y$

f) $\dot{x} = x+5y$
$\dot{y} = 12x-y$

g) $\dot{x} = x-3y$
$\dot{y} = 6x-2y$.

2. Pruébese que el estado de equilibrio del sistema

$$\begin{aligned}\dot{x} &= a_{11}x+a_{12}y\\ \dot{y} &= a_{21}x+a_{22}y\end{aligned}$$

es estable si

$$a_{11}a_{22}-a_{12}a_{21} > 0 \quad \text{y} \quad a_{11}+a_{22} < 0;$$

inestable si

$$a_{11}a_{22}-a_{12}a_{21} < 0 \quad \text{o} \quad a_{11}+a_{22} > 0.$$

3. Pruébese que:

a) Si $\mathbf{u}+i\mathbf{v}$ y $\mathbf{u}-i\mathbf{v}$ son vectores linealmente independientes, donde $\mathbf{u}$ y $\mathbf{v}$ son vectores reales, entonces $\mathbf{u}$ y $\mathbf{v}$ son linealmente independientes.

b) Si A es una matriz real con valores característicos complejos, entonces las partes real e imaginaria de un vector característico son vectores linealmente independientes (es decir, si $\mathbf{w} = \mathbf{u}+i\mathbf{v}$ es un vector característico de A, con $\mathbf{u}$ y $\mathbf{v}$ reales, entonces $\mathbf{u}$ y $\mathbf{v}$ son linealmente independientes).

4. Determínese la solución de estado estable de

a) $\ddot{x}+3\dot{x}+25x = \cos \omega t$

b) $\ddot{x}+0.003\dot{x}+25x = 1+\cos \omega t$

c) $\dot{x} = x+2y+\cos t$
$\dot{y} = -3x-2y+\operatorname{sen} t$.

5. Determínese en qué frecuencia angular ω la amplitud de la solución de estado estable toma su valor máximo y determínese la amplitud máxima.

a) $\dot{x}+5x = \cos \omega t$

b) $10\ddot{x}+25\dot{x}+160x = \cos \omega t$

c) $10\ddot{x}+0.001\dot{x}+160x = \cos \omega t$

d) $10\ddot{x}+0.001\dot{x}+160x = \cos(\omega t+\delta)$.

6. Determínese la solución de estado estable de $\ddot{x}+\dot{x}+2x = f$ donde f es periódica de periodo 1 y

$$f(t) = \begin{cases} 1, & 0 \leqslant t < \delta \\ 0, & \delta \leqslant t < 1 . \end{cases}$$

7. Pruébese que si $\mathbf{f}$ es periódica de periodo T y si $\dot{\mathbf{x}} = A\mathbf{x}+\mathbf{f}$ tiene una solución acotada para todo $t \geqslant 0$, entonces el sistema tiene una solución periódica de periodo T.

11. OSCILACIONES LINEALES. $\ddot{x}+2\alpha\dot{x}+\omega_0{}^2x = f$

El estudio de las oscilaciones (vibraciones) de los sistemas mecánicos y eléctricos siempre envuelve una idealización considerable del sistema físico, y una aproximación primera al sistema bastante natural es suponer que éste es lineal. El comportamiento de los sistemas simples (un grado de libertad) puede a menudo describirse por una ecuación diferencial de la forma

11.1 $$\ddot{x}+2\alpha\dot{x}+\omega_0{}^2x = f$$

donde α y ω_0 son constantes positivas. Un ejemplo mecánico típico es el de una masa m unida a un muelle "elástico" con frenamiento "viscoso" (figura 13). Los términos "elástico" y "viscoso" indican que el muelle y

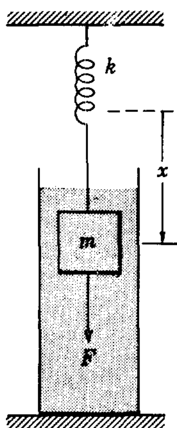


FIGURA 13

el frenamiento se supone que son lineales. Denotemos por $x(t)$ el desplazamiento de la masa m desde su posición de equilibrio a su posición en el instante t. La fuerza del muelle es $-kx(t)$ y la fuerza de freno es $-\beta\dot{x}(t)$. La constante k mide la rigidez del muelle y se llama "constante del muelle" y a β se llama "constante de freno". Sea $F(t)$ la fuerza aplicada (externa) en el instante t. La ecuación vectorial de movimiento está de acuerdo con la segunda ley de Newton

$$m\ddot{x}(t) = -\beta\dot{x}(t)-kx(t)+F(t)$$

o bien

11.2 $$m\ddot{x}+\beta\dot{x}+kx = F.$$

Una analogía eléctrica de este oscilador mecánico simple es un circuito con una resistencia, una bobina y un condensador en serie (figura 14). Sea $q(t)$ la carga del condensador, $i(t) = \dot{q}(t)$ es la corriente, y $E(t)$ es el

voltaje aplicado. Los tres elementos están idealizados como elementos del circuito linealmente colocados. La caída de voltaje a través de la bobina es $LD_t i(t) = L\ddot{q}(t)$, y la caída de voltaje a través del condensador es $\frac{1}{C}q(t)$. Por la segunda ley de Kirchhoff

11.3 $$L\ddot{q}+R\dot{q}+\frac{1}{C}q = E.$$

A L se le llama "inductancia", a R "resistencia", y a C "capacitancia".

El comportamiento de cada uno de los anteriores sistemas está descrito por una ecuación diferencial de la forma 11.1 con coeficientes positivos. En el ejemplo mecánico:

$$2\alpha = \beta/m, \quad \omega_0{}^2 = k/m, \quad \text{y } f = F/m.$$

En el ejemplo eléctrico

$$2\alpha = R/L, \quad \omega_0{}^2 = 1/LC, \quad \text{y } f = E/L.$$

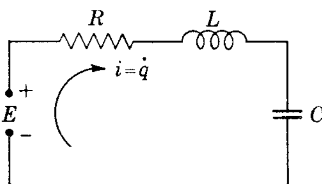


FIGURA 14

Movimientos libres. Suponemos aquí que no hay fuerzas (o voltajes) externas: $f = 0$. La ecuación diferencial es entonces la ecuación homogénea

11.4 $$\ddot{x}+2\alpha\dot{x}+\omega_0{}^2 x = 0.$$

Consideremos primero el caso en que no existe frenaje alguno: $\alpha = 0$. Entonces la ecuación diferencial es

$$\ddot{x}+\omega_0{}^2 x = 0,$$

cuya solución general es

$$x(t) = a\cos(\omega_0 t+\delta).$$

La frecuencia de la oscilación es ω_0 y se le llama *frecuencia natural de oscilación* del sistema. De donde para los sistemas mecánicos y eléctricos la frecuencia natural de oscilación está dada por

11.5 $$\omega_0 = \sqrt{\frac{k}{m}} \quad \text{y} \quad \omega_0 = \frac{1}{\sqrt{LC}}.$$

La frecuencia natural ω_0 no depende de las condiciones iniciales y es una característica importante del sistema. Nos daremos más cuenta de la importancia de esto a medida que nuestra discusión progrese. La *amplitud A* y la *fase* δ de la oscilación libre depende de las condiciones iniciales.

Con frenaje ($\alpha > 0$) la ecuación característica para las oscilaciones libres es

11.6 $$\varphi(\lambda) = \lambda^2+2\alpha\lambda+\omega_0{}^2 = 0$$

donde ($\alpha = \beta/2m$, $\alpha = R/2L$).

Las raíces características son

11.7 $$\lambda_1 = -\alpha+\sqrt{\alpha^2-\omega_0{}^2}, \quad \lambda_2 = -\alpha-\sqrt{\alpha^2-\omega_0{}^2}.$$

Queremos discutir las curvas solución en el plano así como las gráficas de x, y para esto introduciremos el sistema equivalente

11.8 $$\begin{aligned}\dot{x} &= y\\ \dot{y} &= -\omega_0{}^2x-2\alpha y.\end{aligned}$$

Aquí $A = \begin{pmatrix} 0 & 1 \\ -\omega_0{}^2 & -2\alpha \end{pmatrix}$. Los valores característicos de A están dados por 11.7 y vectores característicos correspondientes son $\mathbf{v}^1 = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix}$, $\mathbf{v}^2 = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}$. En el ejemplo mecánico x es el desplazamiento y y la velocidad. En el ejemplo eléctrico x es la carga y y la corriente. El origen ($x = 0$, $y = 0$) es el único punto de equilibrio del sistema bajo nuestra hipótesis de que $\alpha > 0$ y $\omega_0{}^2 > 0$, que se aplica a nuestros ejemplos físicos. Las raíces características tienen partes reales negativas. El estado de equilibrio $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ es asintóticamente estable: *con frenaje todo movimiento libre tiende al origen (el estado de equilibrio) cuando t tiende a infinito* ($x(t)\to 0$ *y* $\dot{x}(t)\to 0$ *cuando* $t\to\infty$).

Caso 1. Asentamiento. $\alpha^2 > \omega_0{}^2 (\beta^2 > 4km,\ R^2C > 4L)$.

Las raíces características λ_1, λ_2 son reales y negativas ($\lambda_2 < \lambda_1 < 0$) y los vectores característicos son reales. La solución general es

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 e^{\lambda_1 t}\mathbf{v}^1 + c_2 e^{\lambda_2 t}\mathbf{v}^2;$$

es decir,

$$x(t) = c_1 e^{\lambda_1 t}+c_2 e^{\lambda_2 t}$$

$$y(t) = \dot{x}(t) = c_1\lambda_1 e^{\lambda_1 t}+c_2\lambda_2 e^{\lambda_2 t}.$$

El origen es un nodo estable. Las curvas solución se muestran en la

figura 16. Para el circuito eléctrico generalmente se está más interesado en la corriente (es más facil de medir) y se puede graficar $y = \dot{x}$ en lugar de x. La figura 15 muestra el comportamiento tanto de x como de $\dot{x}$.

Caso 2. Oscilaciones amortiguadas $\alpha^2 < \omega_0{}^2 (\beta^2 < 4km,\ R^2 C < 4L)$.

Las raíces características son complejas con partes reales negativas. El vector característico $\mathbf{v}^1$ es

$$\mathbf{v}^1 = \begin{pmatrix} 1 \\ -\alpha + i\omega_1 \end{pmatrix} \text{ donde } \omega_1{}^2 = \omega_0{}^2 - \alpha^2.$$

De aquí que $\mathbf{v}^1 = \mathbf{u} + i\mathbf{v}$, donde $\mathbf{u} = \begin{pmatrix} 1 \\ -\alpha \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} 0 \\ \omega_1 \end{pmatrix}$, y la solución general es (10.4, pág. 647)

$$\begin{pmatrix} x \\ y \end{pmatrix} = ae^{-\alpha t}[\cos(\omega_1 t + \delta)\mathbf{u} - \operatorname{sen}(\omega_1 t + \delta)\mathbf{v}];$$

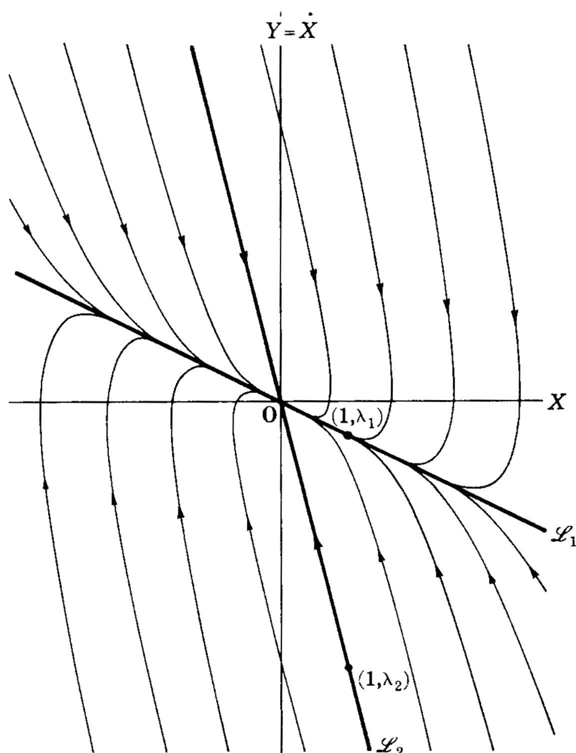
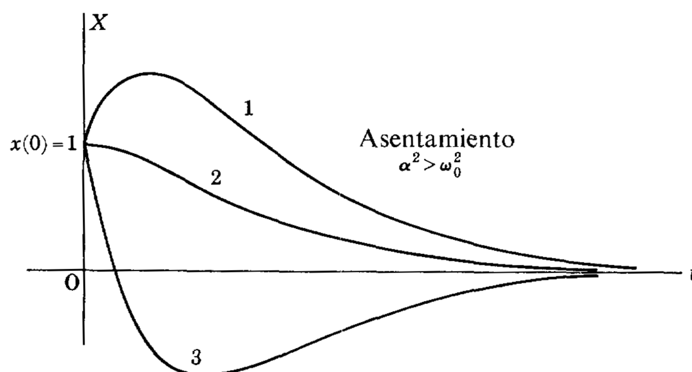


FIGURA 15 $\alpha^2 > \omega_0{}^2$



1. Estado inicial en el primer cuadrante (figura 15):
2. estado inicial sobre el eje X;
3. estado inicial en el cuarto cuadrante debajo de $\mathscr{L}_2$

FIGURA 16

es decir,

$$x(t) = ae^{-\alpha t} \cos(\omega_1 t + \delta)$$
$$\dot{x}(t) = y(t) = -ae^{-\alpha t}[\alpha \cos(\omega_1 t + \delta) + \omega_1 \operatorname{sen}(\omega_1 t + \delta)].$$

El origen es un foco estable (figura 17) y tenemos oscilaciones amortiguadas (figura 18).

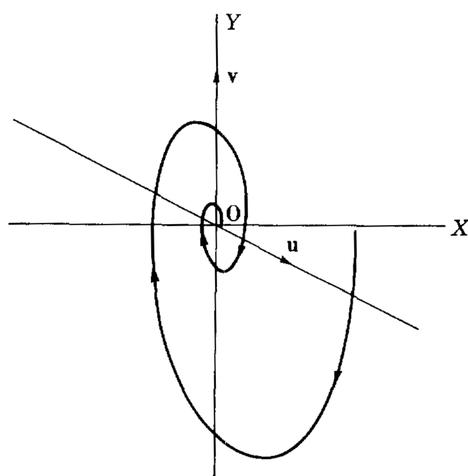


FIGURA 17 $\alpha^2 < \omega_0^2$

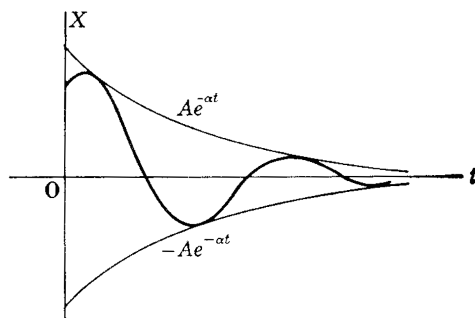


FIGURA 18 Oscilación amortiguada $\alpha^2 < \omega_0{}^2$

Caso 3. Frenaje crítico.

Es éste el caso frontera entre el asentamiento y las oscilaciones amortiguadas. Aquí $\lambda_1 = \lambda_2 = -\alpha$. Hay una sola dirección característica

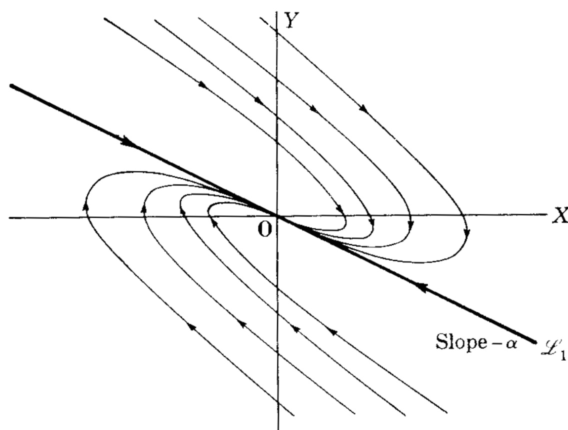


FIGURA 19 $\alpha^2 = \omega_0{}^2$

$\mathbf{v}^1 = \begin{pmatrix} 1 \\ -\alpha \end{pmatrix}$. La solución general es (pág. 612)

$$\mathbf{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{-\alpha t}\left[(c_1 + c_2 t)\mathbf{v}^1 + c_2\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right];$$

es decir

$$x(t) = (c_1 + c_2 t)e^{-\alpha t}$$

$$y(t) = (c_2 - \alpha c_1 - \alpha c_2 t)e^{-\alpha t}.$$

El origen es un nodo estable (figura 19).

Oscilaciones forzadas. Sabemos por lo que hemos aprendido en la sección precedente que, como las raíces características tienen partes reales negativas, la ecuación

$$\ddot{x}+2\alpha\dot{x}+\omega_0^2 x = f$$

tiene una solución de estado estable para cada término forzante periódico f. Toda solución entonces tiende según que $t\to\infty$ a esta solución de estado estable. Es particularmente fácil determinar la solución de estado estable para entradas sinusoidales cos ωt. Consideramos las ecuaciones

11.9 $$c(\ddot{x}+2\alpha\dot{x}+\omega_0^2 x) = a\cos\omega t$$

y

11.10 $$c(\ddot{x}+2\alpha\dot{x}+\omega_0^2 x) = ae^{i\omega t}, \quad \text{donde } c = m \text{ o } L.$$

La parte real de la solución de estado estable de 11.10 es la solución de estado estable de 11.9. Busquemos una solución de estado estable de 11.10 de la forma

$$x(t) = he^{i\omega t}.$$

Sustituyendo en 11.10 obtenemos como la ecuación para h

$$c((i\omega)^2+2\alpha(i\omega)+\omega_0 2)he^{i\omega t} = ae^{i\omega t}$$

o bien

$$h = \frac{a}{\varphi(i\omega)},$$

donde $\varphi(\lambda) = c(\lambda^2+2\alpha\lambda+\omega_0^2)$ es el polinomio característico. De donde la solución de estado estable de 11.10 es

$$x(t) = \frac{a}{\varphi(i\omega)}e^{i\omega t}.$$

Haciendo $\varphi(i\omega) = \rho(\omega)e^{i\delta(\omega)}$, obtenemos como solución de estado estable de 11.9

$$x(t) = \mathrm{Rl}\left[\frac{a}{\varphi(i\omega)}e^{i\omega t}\right] = \mathrm{Rl}\left[\frac{a}{\rho(\omega)}e^{i(\omega t-\delta(\omega))}\right]$$

$$= \frac{a}{\rho(\omega)}\cos(\omega t-\delta(\omega)),$$

donde

$$\rho^2(\omega) = |\varphi(i\omega)|^2 = [4\alpha^2\omega^2+(\omega^2-\omega_0^2)^2]c^2$$

y

$$\delta(\omega) = \arg(\varphi(i\omega)) = \tfrac{1}{2}\pi + \arg[2\alpha\omega+i(\omega^2-\omega_0^2)].$$

La solución de estado estable tiene el mismo periodo que el término forzante. Su amplitud es la de la entrada amplificada por el factor $k(\omega) = 1/\rho(\omega)$, y su fase está corrida respecto a la de la entrada en $\delta(\omega)$ radianes. Ahora

11.11 $$\varphi(i\omega) = i\omega\left[2\alpha + \frac{i}{\omega}(\omega^2 - \omega_0{}^2)\right]c.$$

En el ejemplo mecánico $c = m$, $\alpha = \beta/2m$, $\omega_0{}^2 = k/m$, y

$$\varphi(i\omega) = i\omega\left[\beta + i\left(\omega m - \frac{k}{\omega}\right)\right].$$

En el ejemplo eléctrico $c = L$, $\alpha = R/2L$, $\omega_0{}^2 = 1/LC$, y

$$\varphi(i\omega) = i\omega\left[R + i\left(\omega L - \frac{1}{\omega C}\right)\right].$$

Estudiaremos primero el factor de amplificación $k(\omega) = 1/\rho(\omega)$. Encontremos dónde es un máximo $k(\omega)$, que es donde $\rho(\omega)$ es un mínimo. Nótese que

$$D_{\omega^2}[\rho^2(\omega)] = 2c^2(2\alpha^2 + \omega^2 - \omega_0{}^2).$$

Caso 1. $\omega_0{}^2 \leqslant 2\alpha^2$. Entonces $\omega^2 + 2\alpha^2 - \omega_0{}^2 > 0$ para $\omega > 0$: $\rho(\omega)$ aumenta con ω y el mínimo de $\rho(\omega)$ ocurre en $\omega = 0$. De donde $k_{\text{máx}} = 1/\rho(0) = 1/c\omega_0{}^2$ (figura 20).

Caso 2. $\omega_0{}^2 > 2\alpha^2$. El mínimo de $\rho(\omega)$ ocurre en $\omega_1{}^2 = \omega_0{}^2 - 2\alpha^2$, y $k_{\text{máx}} = \dfrac{1}{2\alpha c\sqrt{\omega_0{}^2 - \alpha^2}}$ (figura 21). Nótese que cuando $\alpha \to 0$, $\omega_1 = \sqrt{\omega_0{}^2 - 2\alpha^2} \to \omega_0$ y $k_{\text{máx}} \to \infty$. El máximo de $k(\omega)$ ocurre en la frecuencia ω_1, que es siempre menor que la frecuencia de resonancia. (Ésta corresponde a la amplitud del desplazamiento de estado estable y la carga del estado estable.)

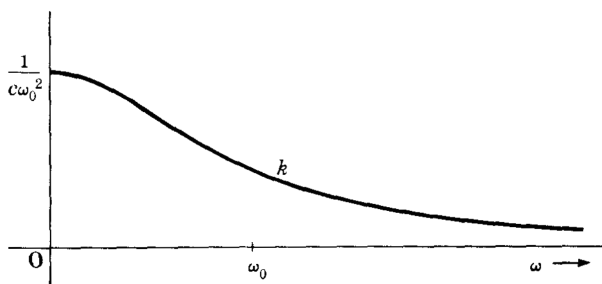


FIGURA 20 Amplificación (desplazamiento, carga) y frecuencia

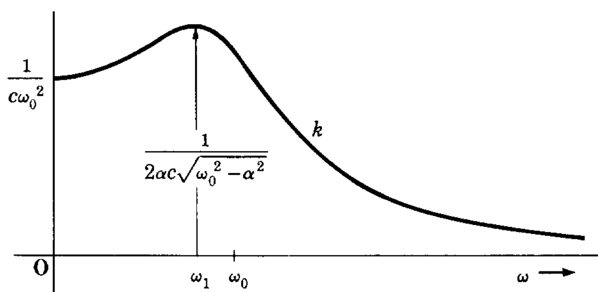


FIGURA 21 Amplificación (desplazamiento, carga) y frecuencia ($\omega_0{}^2 \leqslant 2\alpha^2$)

El corrimiento de la fase $\delta(\omega) = \frac{1}{2}\pi + \arg\left[2\alpha + \frac{i}{\omega}(\omega^2 - \omega_0{}^2)\right]$ se muestra en la figura 22.

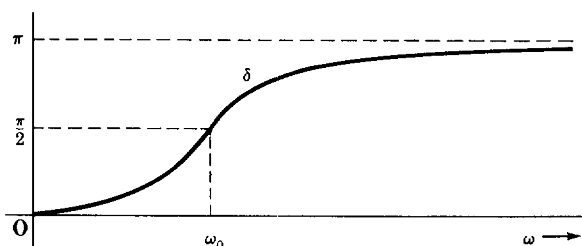


FIGURA 22 Corrimiento de fase (desplazamiento, carga) y frecuencia ($\omega_0{}^2 > 2\alpha^2$)

Cuando se discute el circuito eléctrico, se está más interesado, en general, como antes señalábamos, en la corriente. La corriente corresponde a $\dot{x}$, y el estado estable $\dot{x}$ es

$$\dot{x}(t) = \mathrm{Rl}\left[\frac{i\omega}{\varphi(i\omega)}\, ae^{i\omega t}\right].$$

La cantidad $z(\omega) = \varphi(i\omega)/i\omega$ se llama *impedancia* del circuito. Entonces el estado estable $\dot{x}$ es

$$\dot{x}(t) = \mathrm{Rl}\left[\frac{a}{z(\omega)}\, e^{i\omega t}\right]$$

donde $z(\omega) = \left[2\alpha + \frac{i}{\omega}(\omega^2 - \omega_0{}^2)\right]c$. Para el oscilador mecánico

$$z(\omega) = \beta + i\left(\omega m - \frac{k}{\omega}\right)$$

y para el oscilador eléctrico

$$z(\omega) = R + i\left(\omega L - \frac{1}{\omega C}\right).$$

La impedancia de una resistencia es R, la impedancia de una inductancia es $i\omega L$, y la impedancia de un condensador es $1/i\omega C$. La impedancia de los elementos en serie es la suma de las impedancias.

Haciendo $z(\omega) = \gamma(\omega)e^{i\theta(\omega)}$, tenemos para el estado estable $\dot{x}$ (velocidad o corriente)

$$\dot{x}(t) = \frac{a}{\gamma(\omega)} \cos(\omega t - \theta(\omega)),$$

donde $\gamma^2(\omega) = |z(\omega)|^2 = \left[4\alpha^2 + \left(\omega - \frac{\omega_0^2}{\omega}\right)^2\right]c^2$ y $\theta(\omega) = \arg(z(\omega)) = [2\alpha\omega + i(\omega^2 - \omega_0^2)]$.

El mínimo de $\gamma(\omega)$ siempre ocurre en resonancia ($\omega = \omega_0$) y, por tanto, el máximo del factor de amplificación $1/\gamma(\omega)$ ocurre en $\omega = \omega_0$. El factor de amplificación $1/\gamma(\omega)$ y el corrimiento de fase $\theta(\omega)$ se muestran en las figuras 23 y 24.

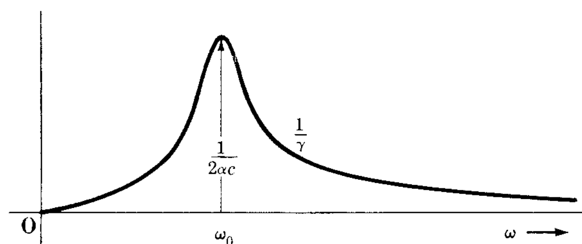


FIGURA 23 Amplificación (velocidad, corriente) y frecuencia

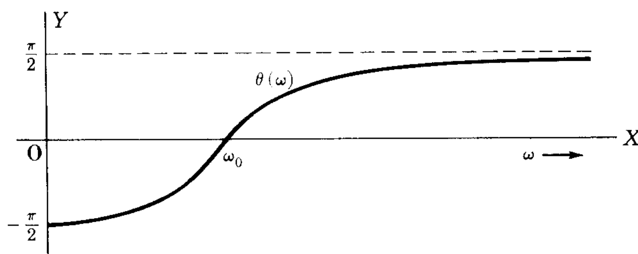


FIGURA 24 Corrimiento de fase (velocidad, corriente) y frecuencia

11.12 Ejemplo. Encuéntrese la solución del estado estable de

$$0.5\ddot{x} + 0.1\dot{x} + 50x = 1 + \operatorname{sen}^3 \omega t$$

en $\omega = \omega_0 = 10$.

SOLUCIÓN.

$$\text{sen}^3\,\omega t = \frac{i}{2^3}(e^{i\omega t}-e^{-i\omega t})^3$$

$$= \frac{i}{8}(e^{3i\omega t}-3e^{i\omega t}+3e^{-i\omega t}-e^{-3i\omega t})$$

$$= 2\text{Rl}\left[\frac{i}{8}(e^{3i\omega t}-3e^{i\omega t})\right].$$

Queremos, pues, resolver

11.13 $$0.5\ddot{x}+0.1\dot{x}+50x = \text{Rl}\left[1-\frac{i}{4}(3e^{i\omega t}-e^{3i\omega t})\right].$$

Aquí

$$\varphi(i\omega) = -0.5\omega^2+0.1\,i\omega+50,$$

y la solución de estado estable de

$$0.5\ddot{x}+0.1\dot{x}+50x = ae^{i\omega t}$$

es

$$\frac{a}{\varphi(i\omega)}e^{i\omega t}.$$

De donde, de acuerdo con el principio de superposición, la solución de estado estable de 11.13 es

$$x(t) = \text{Rl}\left[\frac{1}{\varphi(0)}-\frac{i}{4}\left(\frac{3}{\varphi(i\omega)}e^{i\omega t}-\frac{1}{\varphi(3i\omega)}e^{3i\omega t}\right)\right].$$

En particular, queremos la solución de estado estable correspondiente a $\omega = 10$. Obtenemos, entonces,

$$x(t) = \frac{1}{50}-\text{Rl}\left[\frac{i}{4}\left(\frac{3}{i}e^{i10t}-\frac{1}{-400+3i}e^{i30t}\right)\right]$$

$$= \frac{1}{50}-\frac{3}{4}\cos 10t+\frac{1}{4(16\times 10^4+9)}(3\cos 30t+400\,\text{sen}\,30t).$$

De donde $x(t) = \frac{1}{50}-\frac{3}{4}\cos 10t$ es una buena aproximación de la solución de estado estable.

Problemas

1. Dibújese en el plano $(x, \dot{x})$ (el plano fase) las curvas solución de

a) $\ddot{x}+5\dot{x}+6x = 0$
b) $\ddot{x}+2\dot{x}+5x = 0$
c) $\ddot{x}+2\dot{x}+x = 0$.

2. El sistema $\ddot{x}+5\dot{x}+6x=0$ comienza en el instante $t=0$ con $x(0)=1$ y $\dot{x}(0)=v$. ¿Para qué valores de v permanece positivo $x(t)$ para:

a) todo $t>0$?
b) todo t?

3. Determínese la solución de estado estable para x y $\dot{x}$ de

a) $10^3\ddot{x}+5\dot{x}+10^5 x = \cos 3t$
b) $10^3\ddot{x}+5\dot{x}+10^5 x = \operatorname{sen} 3t$
c) $\ddot{x}+16\dot{x}+36x = a\cos\omega t$
d) $\ddot{x}+16\dot{x}+36x = a\cos(\omega t+\delta)$
e) $\ddot{x}+0.03\dot{x}+100x = \operatorname{sen} 10t+\operatorname{sen}^3 10t$
f) $\ddot{x}+3\dot{x}+2x = |\cos t|$
g) $x+7\ddot{x}+14\dot{x}+8x = \operatorname{sen}\omega t$
h) $\dfrac{d^4x}{dt^4} + 6\dfrac{d^3x}{dt^3} + 13\dfrac{d^2x}{dt^2} + 12\dfrac{dx}{dt} + 4x = \cos\omega t.$

4. Determínese la solución general de

a) $\ddot{x}+4\dot{x}+4x = e^{-t}\cos 3t$
b) $\ddot{x}+\omega_0{}^2 x = e^{-2\alpha t}\cos\omega_0 t$
c) $\ddot{x}+3\dot{x}-10x = a\cos\omega t$.

5. Determínese la amplitud de la corriente de estado estable $i=\dot{q}$ para el circuito cuya ecuación diferencial es

a) $0.10\ddot{q}+100\dot{q}+2\times 10^4 q = E\cos 120\pi t$
b) $0.10\ddot{q}+0.10\dot{q}+14\times 10^4 q = E\cos 120\pi t$
c) $0.10\ddot{q}+10^5\dot{q}+14\times 10^4 q = E\cos 120\pi t$.

6. Una boya cilíndrica de a centímetros de diámetro, pesa W kilogramos. Flota en una posición vertical en agua cuya densidad es α gramos/cm^3. Proporciónese una fórmula para su periodo natural de oscilación.

7. En muchos instrumentos sencillos de medida el impulso de entrada a medir es $f(t)=c$ para $t\geqslant 0$, y la lectura $\omega_0{}^2 y$ del instrumento satisface la ecuación diferencial $(t\geqslant 0)$

$$\ddot{y}+2\alpha\dot{y}+\omega_0{}^2 y = c, \quad y(0)=\dot{y}(0)=0.$$

Demuéstrese que si $\alpha>0$, entonces $\omega_0{}^2 y(t)\to c$ cuando $t\to\infty$. El error $x(t)$

en un momento cualquiera $t \geqslant 0$ es entonces $x(t) = c - \omega_0{}^2 y(t)$. Para $t \geqslant 0$

$$\ddot{x} + 2\alpha\dot{x} + \omega_0{}^2 x = 0, \qquad x(0) = c, \quad \dot{x}(0) = 0$$

Una medida de cuán rápidamente el error tiende a cero es $\int_0^\infty |x(t)|^2\, dt$. Determínese el valor de α que para todo c minimiza $\int_0^\infty |x(t)|^2\, dt$.

8. Con la notación del problema 7 supongamos que $x(0) = 0$, $\dot{x}(0) = v$. Determínese el valor de α que minimiza $\int_0^\infty |x(t)|^2\, dt$ para todo v.

12. ECUACIONES EXACTAS

En toda esta sección $F(x, y)$ representará una función real con derivadas parciales de primer orden continuas sobre un conjunto abierto $\mathscr{E}$ del plano R^2. $A(x, y)$ y $B(x, y)$ serán un par de funciones reales continuas sobre $\mathscr{E}$. Usamos γ para denotar una curva diferenciable en $\mathscr{E}$ cuya parametrización es $x = u(t)$, $y = v(t)$, t en $\langle a, b\rangle$. La ecuación diferencial

12.1 $$A(x, y)\dot{x} + B(x, y)\dot{y} = 0$$

es entonces una ecuación diferencial de curvas γ con la propiedad de que en cada punto (x, y) la curva γ es ortogonal a $(A(x, y), B(x, y))$. Estrechamente asociada con esta ecuación diferencial tenemos la *forma diferencial* $A(x, y)dx + B(x, y)dy$. Nuestro primer objetivo es ver cuál es la relación entre la forma diferencial $A(x, y)dx + B(x, y)dy$ y la ecuación diferencial 12.1.

Recuérdese (pág. 000) que la forma diferencial $A(x, y)dx + B(x, y)dy$ se dice que es exacta sobre $\mathscr{E}$ si hay una función F con la propiedad de que

12.2 $$dF(x, y;\, dx, dy) = A(x, y)dx + B(x, y)dy$$

para todo (x, y) en $\mathscr{E}$ y cualesquier dx, dy. En lugar de 12.2 escribimos

$$dF = A(x, y)dx + B(x, y)dy.$$

Según la definición de la diferencial, se sigue entonces que $A(x, y)dx + B(x, y)dy$ es exacta sobre $\mathscr{E}$ si y sólo si hay una función F que sobre $\mathscr{E}$ es una solución de

12.3 $$\frac{\partial F}{\partial x} = A(x, y), \quad \frac{\partial F}{\partial y} = B(x, y).$$

12.4 Ejemplo. Determínese si las siguientes formas diferenciales son o

no exactas.

a) $\cos y\,dx - x \operatorname{sen} y\,dy$
b) $(y+x^2)dx + x\,dy$
c) $x\,dy - y\,dx$.

SOLUCIÓN. *a*) Supongamos que $dF = \cos y\,dx - x \operatorname{sen} y\,dy$. Entonces,

$$\frac{\partial F}{\partial x} = \cos y \quad \text{y} \quad \frac{\partial F}{\partial y} = -x \operatorname{sen} y\,.$$

La primera de estas ecuaciones implica que $F(x, y) = x \cos y + u(y)$ para un $u(y)$ arbitrario. Entonces $\dfrac{\partial F}{\partial y} = -x \operatorname{sen} y + u'(y)$ y vemos, tomando $u = 0$, que $dF = d(x \cos y) = \cos y\,dx - x \operatorname{sen} y\,dy$. La forma diferencial es exacta sobre R^2.

b) Observemos que $d(xy) = y\,dx + x\,dy$ y que $d(\frac{1}{3}x^3) = x^2\,dx$. De donde $d(xy + \frac{1}{3}x^3) = (y+x^2)dx + x\,dy$, y la forma diferencial es exacta sobre R^2.

c) Supongamos que $x\,dy - y\,dx$ es exacta con $dF = x\,dy - y\,dx$. Entonces

$$\frac{\partial F}{\partial y} = x \quad \text{y} \quad \frac{\partial F}{\partial x} = -y\,.$$

La primera de estas funciones implica

$$F(x, y) = xy + u(x)$$

para alguna función u. La segunda ecuación implica que

$$\frac{\partial F}{\partial x} = y + u'(x) = -y\,.$$

Esto no puede verificarse sobre ningún conjunto abierto de R^2 y, por tanto, la forma diferencial no es exacta sobre ningún conjunto abierto de R^2.

Si $dF = A(x, y)dx + B(x, y)dy$ sobre $\mathscr{E}$, diremos que $F(x, y) = c$ es una *solución* sobre $\mathscr{E}$ de la ecuación diferencial

12.5 $$A(x, y)dx + B(x, y)dy = 0\,.$$

Así pues, las soluciones sobre $\mathscr{E}$ de 12.5 son las curvas de nivel $F(x, y) = c$ en $\mathscr{E}$. Por ejemplo, las soluciones sobre R^2 de $x\,dx + y\,dy = 0$ son los círculos de la familia de círculos $x^2 + y^2 = c^2$.

Vimos, en el ejemplo 12.4*c*, que $x\,dy - y\,dx$ no es exacta. Sin embargo, el dividir esta forma diferencial por x^2 la hace exacta ($x \neq 0$):

$$d\left(\frac{y}{x}\right) = \frac{x\,dy - y\,dx}{x^2}\,.$$

Si hay una función $\mu(x, y)$ que es continua y no se anula sobre $\mathscr{E}$ y tal que $\mu(x, y)\, A(x, y)dx + \mu(x, y)\, B(x, y)dy$ es exacta sobre $\mathscr{E}$ con $dF = \mu A\, dx + \mu B dy$, diremos también de nuevo que $F(x, y) = c$ es una *solución* sobre $\mathscr{E}$ de 12.5. La función μ se llama *factor de integración*. Todo esto parece puro formulismo, pero ahora veremos por qué está justificado este lenguaje.

Una solución $x = u(t)$, $y = v(t)$ de

12.6 $$A(x, y)\dot{x} + B(x, y)\dot{y} = 0$$

define una curva γ. Hay, por tanto, una sencilla relación entre las curvas solución γ y las curvas de nivel $F(x, y) = c$ que son soluciones de 12.5.

12.7 Teorema. *Sea $F(x, y) = c$ una solución sobre $\mathscr{E}$ de*

$$A(x, y)dx + B(x, y)dy = 0.$$

Si γ es una curva solución en $\mathscr{E}$ de $A(x, y)\dot{x} + B(x, y)\dot{y} = 0$, entonces se encuentra sobre una curva de nivel $F(x, y) = c$. Recíprocamente, si γ es una curva diferenciable que se encuentra sobre una curva de nivel $F(x, y) = c$, entonces γ es una curva solución de $A(x, y)\dot{x} + B(x, y)\dot{y} = 0$.

PRUEBA. Como $dF = \mu(x, y)\, \{A(x, y)dx + B(x, y)dy\}$,

$$\frac{dF}{dt} = D_t[F(u(t), v(t))] = \frac{\partial F}{\partial x} u'(t) + \frac{\partial F}{\partial y} v'(t)$$

$$= \mu(u(t), v(t))\, \{A(u(t), v(t))\, u'(t) + B(u(t), v(t))\, v'(t)\}.$$

Bajo la hipótesis de que la curva γ: $x = u(t)$, $y = v(t)$, $t \in \langle a, b \rangle$, es una curva solución de $A(x, y)\dot{x} + B(x, y)\dot{y} = 0$, se sigue que $\dfrac{dF}{dt} = 0$ sobre $\langle a, b \rangle$. De donde, para alguna cierta constante c, $F(u(t), v(t)) = c$ para todo $t \in \langle a, b \rangle$; es decir, γ está sobre la curva de nivel $F(x, y) = c$. Recíprocamente, γ sobre $F(x, y) = c$ implica $\dfrac{dF}{dt} = 0$. Como $\mu(x, y)$ es un factor de integración sobre $\mathscr{E}$, $\mu(x, y)$ no se anula sobre $\mathscr{E}$; de donde $\dfrac{dF}{dt} = 0$ implica

$$A(u(t), v(t))u'(t) + B(u(t), v(t))\, v'(t) = 0.$$

La curva γ es una curva solución de $A(x, y)\dot{x} + B(x, y)\dot{y} = 0$. Y esto completa la prueba.

El anterior teorema nos dice que si $F(x, y) = c$ es una solución sobre $\mathscr{E}$ de

$$A(x, y)dx + B(x, y)dy = 0,$$

entonces todas las soluciones de $A(x, y)\dot{x} + B(x, y)\dot{y} = 0$ están dadas implícitamente por $F(x, y) = c$. En relación con esto señalamos que como una consecuencia del teorema general de existencia para las ecuaciones diferenciales, resulta que para cada (x_0, y_0) de $\mathscr{E}$ hay funciones u y v diferenciables sobre alguna vecindad $\mathscr{N} = \langle -\delta, \delta \rangle$ tales que $F(u(t), v(t)) = F(x_0, y_0)$ para todo t en $\mathscr{N}$. Imponiendo condiciones un poco más fuertes sobre F, este teorema de la función implícita se prueba en el problema 4, pág. 707.

12.8 Ejemplo. Determínense todas las curvas ortogonales a la familia de circunferencias $x^2 + y^2 = c^2$.

SOLUCIÓN. En un punto (x, y) la tangente a la circunferencia es paralela a $(-y, x)$. La ecuación diferencial de las curvas que buscamos es, por tanto,

$$(-y, x) \cdot (\dot{x}, \dot{y}) = x\dot{y} - y\dot{x} = 0.$$

Consideramos entonces la ecuación diferencial

$$x\,dy - y\,dx = 0.$$

Esta ecuación sabemos que no es exacta (ejemplo 12.4c). Sin embargo,

$$d\left(\frac{y}{x}\right) = \frac{x\,dy - y\,dx}{x^2}, \quad x \neq 0$$

y $\dfrac{1}{x^2}$ es un factor de integración. De donde, para $x \neq 0$, $\dfrac{y}{x} = c$ es una solución. Las curvas son rectas que pasan por el origen.

Ecuaciones separables. Sea $\mathscr{R}$ un rectángulo (intervalo) abierto. Si hay un factor de integración μ sobre $\mathscr{R}$ tal que

$$\mu(x, y)\,A(x, y)\,dx + \mu(x, y)\,B(x, y)\,dy = f(x)\,dx + g(y)\,dy,$$

entonces decimos que las ecuaciones 12.5 y 12.6 son *separables*. Es claro que $f(x)\,dx + g(y)\,dy$ es exacta sobre $\mathscr{R}$. Para verlo basta definir para un (x_0, y_0) en $\mathscr{R}$ fijo

$$F(x, y) = \int_{x_0}^{x} f(s)\,ds + \int_{y_0}^{y} g(s)\,ds.$$

Entonces $dF = f(x)\,dx + g(y)\,dy$.

12.9 Ejemplo. Obténganse soluciones de

a) $y\,dx - x\,dy = 0$

b) $x^2(ye^x - y)\,dx + y^2 x^3 \operatorname{sen} y\,dy = 0$.

SOLUCIÓN. *a*) (Compárese con el ejemplo 12.8.) Dividiendo por xy tenemos

$$\frac{dx}{x} - \frac{dy}{y} = 0.$$

De donde $\ln|x| - \ln|y| = k$ para $x \neq 0$, $y \neq 0$. Y podemos escribir esto como $\frac{x}{y} = c$.

b) Dividiendo por $x^2 y$ obtenemos

$$\frac{1}{x}(e^x - 1)\,dx + y \operatorname{sen} y\,dy = 0.$$

En el primer cuadrante ($x > 0$, $y > 0$) tenemos como una solución

$$F(x, y) = \int_1^x s^{-1} e^s\,ds - \ln x + \operatorname{sen} y - y \cos y = c.$$

Una condición necesaria para la exactitud sobre un conjunto abierto. En el capítulo 5, pág. 299, se probó que si $\mathbf{f} = (A, B) \in C^1$ sobre un conjunto abierto $\mathscr{E}$ y hay una función F tal que $\mathbf{D}F = \mathbf{f}$ sobre $\mathscr{E}$, entonces $F \in C^2$ sobre $\mathscr{E}$ y, por tanto,

$$\frac{\partial^2 F}{\partial y\,\partial x} = \frac{\partial^2 F}{\partial x\,\partial y} \quad \text{sobre } \mathscr{E}.$$

como $\frac{\partial F}{\partial x} = A$ y $\frac{\partial F}{\partial y} = B$ sobre $\mathscr{E}$, obtenemos

12.10 $$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x} \quad \text{sobre } \mathscr{E}$$

como una condición necesaria para la exactitud de la forma diferencial $A(x, y)dx + B(x, y)dy$ sobre el conjunto abierto $\mathscr{E}$.

Una condición suficiente para la exactitud sobre un rectángulo. Supongamos que A y B tienen primeras derivadas parciales continuas sobre un rectángulo $\mathscr{R}$ y que la condición 12.10 se satisface en $\mathscr{R}$. Sea (x_0, y_0) un punto cualquiera en $\mathscr{R}$ y definamos

$$F(x, y) = \int_{x_0}^x A(s, y_0)\,ds + \int_{y_0}^y B(x, s)\,ds.$$

Entonces

$$\frac{\partial F}{\partial x} = A(x, y_0) + \int_{y_0}^y \frac{\partial B(x, s)}{\partial x}\,ds \quad \text{sobre } \mathscr{R}.$$

Usando 12.10 tenemos sobre $\mathscr{R}$ que

$$\frac{\partial F}{\partial x} = A(x, y_0) + \int_{y_0}^{y} \frac{\partial A(x, s)}{\partial s} ds = A(x,y_0) + A(x,y) - A(x, y_0) = A(x,y).$$

Análogamente podemos verificar que $\dfrac{\partial F}{\partial y} = B(x, y)$ sobre $\mathscr{R}$. Con lo que hemos demostrado que

12.11 Teorema. *Si A y B tienen primeras derivadas parciales continuas sobre un rectángulo abierto $\mathscr{R}$ y si*

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x} \quad \textit{sobre } \mathscr{R},$$

entonces la forma diferencial $A(x, y)dx + B(x, y)dy$ es exacta sobre $\mathscr{R}$.

12.12 Ejemplo. Resuélvanse

a) $(3x^2y^2 + 4y + 1)dx + (2x^3y + 4x)dy = 0$

b) $3 \cos y dx - x \text{ sen } y dy = 0$.

SOLUCIÓN. *a*) $\dfrac{\partial A}{\partial y} = 6x^2y + 4$, $\dfrac{\partial B}{\partial x} = 6x^2y + 4$. La ecuación es exacta sobre R^2. Definamos

$$F(x, y) = \int_0^x A(s, 0)ds + \int_0^y B(x, s)ds = \int_0^x ds + \int_0^y (2x^3s + 4x)ds$$

$$= x + x^3y^2 + 4xy.$$

De donde $x + x^3y^2 + 4xy = c$ es una solución sobre R^2.

b) Esta ecuación no es exacta. Busquemos un factor de integración $\mu(x, y)$. Podemos hacer esto si es que podemos encontrar una solución de $\dfrac{\partial(\mu A)}{\partial y} = \dfrac{\partial(\mu B)}{\partial x}$; es decir, queremos una solución de

$$-3(\text{sen } y)\mu + 3 \cos y \frac{\partial \mu}{\partial y} = -(\text{sen } y)\mu - x \text{ sen } y \frac{\partial \mu}{\partial x}.$$

Vemos que si hacemos $\mu(x, y) = \mu(x)$ (una función de sólo x), entonces

$$\text{sen } y\left[x\frac{d\mu}{dx} - 2\mu\right] = 0.$$

Una solución es $\mu(x)=x^2$, y x^2 es un factor de integración para $x\neq 0$. La forma diferencial $3x^2\cos y dx - x^3$ sen $y dy$ es exacta, y podemos ver inmediatamente que una solución para $x\neq 0$ es

$$x^3 \cos y = c.$$

Ecuaciones de primer orden. $B(x, y)\dfrac{dy}{dx} + A(x, y) = 0.$

Esta ecuación es un caso especial de 12.1. Aquí nos restringiremos a una curva γ que es la gráfica de una función: $x=t$, $y=v(t)$, $t\in\langle a, b\rangle$. Si $dF = A(x, y)dx + B(x, y)dy$ sobre $\mathscr{E}$, entonces la gráfica de cada solución de

12.13
$$B(x, y)\frac{dy}{dx} + A(x, y) = 0$$

en $\mathscr{E}$ está sobre la curva de nivel $F(x, y) = c$. Recíprocamente, si la gráfica de $y = v(x)$ en $\mathscr{E}$ está sobre una curva de nivel $F(x, y) = c$ y v es diferenciable, entonces $v(x)$ es una solución de 12.13. Las soluciones de 12.13 están dadas implícitamente por las curvas de nivel $F(x, y) = c$.

12.14 Ejemplo. Determínense todas las soluciones de

$$\frac{dy}{dx} = y^2.$$

SOLUCIÓN. Consideremos la ecuación diferencial

$$dy - y^2 dx = 0.$$

Esta ecuación es separable para $y\neq 0$. Dividiendo por y^2 tenemos

$$y^{-2}dy - dx = -d(y^{-1}+x)$$

y para $y\neq 0$ las soluciones vienen dadas implícitamente por

$$\frac{1}{y} + x = c,$$

de donde las soluciones para $y\neq 0$ son

$$y = \frac{1}{c-x}.$$

Nótese que para $x<c$, y tiende a infinito cuando $x\to c$. La solución que

satisface $y(x_0) = y_0 \neq 0$ es

$$y = \frac{1}{y_0^{-1} + x_0 - x}.$$

Si $y_0 > 0$, la solución que satisface $y(x_0) = y_0$ está definida para $x < y_0^{-1} + x_0$. Si $y_0 < 0$, la solución está definida para $x > y_0^{-1} + x_0$. Por la ecuación original $\dfrac{dy}{dx} = y^2$ vemos que si $y_0 = 0$, la solución es $y = 0$.

Problemas

1. Resuélvanse

a) $(2x+y)dx+(x+2y)dy = 0$
b) $x^3 dx - y^3 dy = 0$
c) $(y+x^3)dx - x\,dy = 0$
d) $y^3 dx - x^3 dy = 0$
e) $(3x^2y^2 + 3 \operatorname{sen} y)dx + (2x^3y + 3x \cos y)dy = 0$
f) $(y-2)dx + xy\,dy = 0$
g) $x\,dx + \sqrt{a^2 - x^2}\,dy = 0$.

2. Supongamos que $A(x, y)$ y $B(x, y)$ tienen primeras derivadas parciales continuas en un rectángulo $\mathscr{R}$. Supongamos que $p(x)$ y $q(y)$ son continuas para todo x y para todo y. Definamos $P(x) = \int p(x)dx$ y $Q(y) = \int q(y)dy$. Pruébese que

$$\mu(x, y) = e^{P(x)+Q(y)}$$

es un factor de integración de $A(x, y)dx + B(x, y)dy$ sobre $\mathscr{R}$ si y sólo si

$$\frac{\partial A}{\partial y} + q(y)A = \frac{\partial B}{\partial x} + p(x)B \quad \text{sobre } \mathscr{R}.$$

3. Úsese el resultado del problema 2 para resolver las siguientes ecuaciones:

a) $y\,dx - (2x+y)dy = 0$
b) $(p(x)y - q(x))dx + dy = 0$
c) $(y + xy + y^3)dx + (2x + 4y^2)dy = 0$.

4. Considérense dos familias de curvas definidas por $F(x, y) = c$ y $G(x, y) = c$. Si en cada punto de una curva de la familia $F(x, y) = c$ la curva es ortogonal a una curva de la familia $G(x, y) = c$, entonces hablamos de ellas como *familias ortogonales* de curvas.

Si $A(x, y)dx + B(x, y)dy = 0$ es una ecuación diferencial de la familia $F(x, y) = c$, pruébese que $B(x, y)dx - A(x, y)dy = 0$ es una ecuación diferencial de la familia $G(x, y) = c$ ortogonal a la $F(x, y) = c$.

5. Encuéntrese la familia de curvas ortogonales a

a) $x^2-y^2 = c^2$
b) $4x-2y = c$
c) circunferencias por el origen con centros sobre el eje X
d) $\dfrac{x^2}{a^2+c} + \dfrac{y^2}{b^2+c} = 1$
e) elipses con centros en el origen y vértices en $(-1, 0)$ y $(1, 0)$.

6. Encuéntrense las curvas de descenso más rápido sobre la superficie de silla de montar $z = xy$.

13. FORMAS DIFERENCIALES E INTEGRALES LINEALES

El estudio que aquí hemos hecho de ecuaciones exactas y formas diferenciales está estrechamente relacionado con las integrales curvilíneas (sección 8 del capítulo 5), y es nuestro propósito examinar ahora esta conexión.

Sea $\mathscr{E}$ un conjunto abierto y conexo de R^2, y supongamos que A y B son funciones reales continuas sobre $\mathscr{E}$. Sea γ una curva lisa a trozos en $\mathscr{E}$ descrita por $x=u(t)$, $y=v(t)$, $t\in[a, b]$. Con la forma diferencial $\Delta = A(x, y)dx+B(x, y)dy$ y la curva γ asociamos la integral curvilínea

13.1
$$\Delta(\gamma) = \int_\gamma \Delta = \int_\gamma [A\,dx+B\,dy];$$

es decir, definimos,

$$\Delta(\gamma) = \int_a^b [A(u(t), v(t))\,u'(t)+B(u(t), v(t))\,v'(t)]\,dt.$$

13.2 Ejemplo. Evalúese

$$\Delta(\gamma) = \int_\gamma [x^2\,dx+(x^2-y^2)\,dy]$$

a lo largo del arco de parábola γ: $x=t$, $y=t^2$, $t\in[0, 1]$.

SOLUCIÓN. $\Delta(\gamma) = \displaystyle\int_0^1 [t^2+(t^2-t^4)2t]\,dt = \frac{1}{2}.$

Esta asociación de una integral curvilínea con una forma diferencial nos permite dar una interpretación de una forma diferencial Δ. Δ es una función real con ecuación 13.1 como regla de correspondencia y el conjunto de

todas las curvas γ lisas a trozos en $\mathscr{E}$ como dominio. Sabemos (sección 8, capítulo 5) que

$$\Delta(-\gamma) = -\Delta(\gamma).$$

Invirtiendo la orientación de γ cambia el signo de la integral curvilínea. Si $\gamma = \gamma_1 + \gamma_2 + \cdots + \gamma_n$ es la unión de un número finito de curvas lisas γ, entonces, pág. 296,

$$\Delta(\gamma) = \Delta(\gamma_1 + \gamma_2 + \cdots + \gamma_n) = \Delta(\gamma_1) + \Delta(\gamma_2) + \cdots + \Delta(\gamma_n).$$

13.3 Ejemplo. Evalúese

$$\int_\gamma \Delta = \int_\gamma [(x+y^3)\,dx + xy\,dy]$$

a lo largo del cuadrado γ que se muestra en la figura 25.

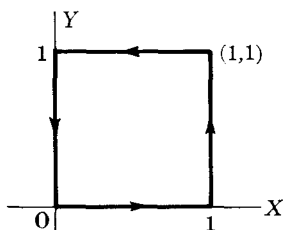


FIGURA 25

SOLUCIÓN. Tomando la parametrización obvia para cada uno de los segmentos rectilíneos, tenemos

$$\int_\gamma \Delta = \int_0^1 x\,dx + \int_0^1 y\,dy + \int_1^0 (x+1)\,dx + \int_1^0 0\,dy = -\tfrac{1}{2}.$$

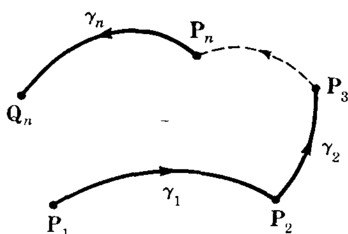


FIGURA 26

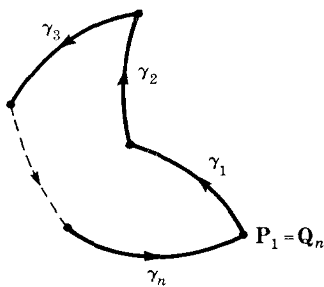


FIGURA 27

Recuérdese que una curva γ se dice que es *lisa a trozos* (figura 26) si $\gamma = \gamma_1 + \gamma_2 + \cdots + \gamma_n$ es la unión de curvas lisas

$$\gamma_i\colon\ x = u_i(t), \quad y = v_i(t), \quad t \in [a_i, b_i], \quad i = 1, \ldots, n,$$

donde el punto terminal $\mathbf{Q}_i = (u_i(b_i), v_i(b_i))$ de γ_1 es el punto inicial $\mathbf{P}_{i+1} = (u_{i+1}(a_{i+1}), v_{i+1}(a_{i+1}))$ de γ_{i+1}, $i = 1, \ldots, n-1$. Si, además, el punto terminal $\mathbf{Q}_n$ de γ_n es el punto inicial $\mathbf{P}_1$ de γ_1, la curva γ se dice que es una curva *cerrada* lisa a trozos (figura 27).

Supongamos ahora que $\Delta = A(x, y)dx + B(x, y)dy$ es exacta sobre $\mathscr{E}$ con $\Delta = dF$. Sea γ una curva lisa a trozos en $\mathscr{E}$ que va del punto $\mathbf{P}$ al punto $\mathbf{Q}$ ($\mathbf{P}$ es el punto inicial de γ y $\mathbf{Q}$ es el punto terminal). Entonces $A\dot{x} + B\dot{y} = \dfrac{dF}{dt}$ la derivada de F a lo largo de γ, y

13.4 $$\int_\gamma \Delta = \int_\gamma dF = \int_a^b \frac{dF}{dt}dt = F(\mathbf{Q}) - F(\mathbf{P}).$$

Decir que Δ es exacta con $\Delta = dF$ es equivalente a decir que $(A, B) = \Delta F$. Este resultado sobre las integrales curvilíneas nos da un medio sencillo de evaluar algunas de ellas. Nos dice también que si Δ es exacta en $\mathscr{E}$ entonces la integral curvilínea $\int_\gamma \Delta$ depende solamente de cuáles sean los puntos inicial y terminal de γ y no de cuál sea la curva particular en $\mathscr{E}$ que una los dos puntos a lo largo de la cual la integral se calcule.

13.5 Ejemplo. Evalúese

$$\Delta(\gamma) = \int_\gamma (y \operatorname{sen} xy\, dx + x \operatorname{sen} xy\, dy)$$

a lo largo de la curva γ: $x = e^{t(t-\pi)} \cos t$, $y = e^{t^2} \operatorname{sen} t$, $t \in [0, \pi]$.

SOLUCIÓN 1. Si nos damos cuenta de que $d(-\cos xy) = y \operatorname{sen} xy dx + x \operatorname{sen} xy dy$, entonces, como la curva va de $(1, 0)$ a $(-1, 0)$ sabemos, puesto que Δ es exacta, que $\Delta(\gamma) = -\cos 0 + \cos 0 = 0$.

SOLUCIÓN 2. $\dfrac{\partial}{\partial y} y \operatorname{sen} xy = \dfrac{\partial}{\partial x} x \operatorname{sen} xy$ en R^2. Por tanto Δ es exacta en R^2 y la integral curvilínea es independiente de la trayectoria. Integrando a lo largo del segmento rectilíneo de $(1, 0)$ a $(-1, 0)$, vemos que, como $y = 0$ a lo largo de esta curva, $\Delta(\gamma) = 0$.

Estos resultados tienen muchas interpretaciones físicas importantes. Sea, como en la sección 9 del capítulo 5, $F(x, y) = (A(x, y), B(x, y))$ una función vectorial continua con un dominio $\mathscr{E}$ en el plano R^2. Supongamos que $\mathbf{F}(\mathbf{x})$ es la fuerza que actúa sobre una partícula unitaria en el punto $\mathbf{x} = (x, y)$. La forma diferencial $\Delta = A(x, y)dx + B(x, y)dy$ puede escribirse entonces como

$$\Delta = \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x},$$

donde $\mathbf{x} = (x, y)$ y $d\mathbf{x} = (dx, dy)$. La integral curvilínea

$$\int_\gamma \Delta = \int_\gamma \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x}$$

es el trabajo efectuado al mover la partícula a lo largo de γ desde su posición inicial hasta el punto terminal. El campo $\mathbf{F}$ se dice que es *conservativo* en $\mathscr{E}$ si $\int_\gamma \Delta = \int_\gamma \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x}$ es cero a lo largo de toda curva cerrada γ contenida en $\mathscr{E}$; es decir, el trabajo efectuado al mover una partícula a lo largo de una curva cerrada es cero. Por curva entendemos aquí una curva lisa a trozos.

Supongamos que el campo $\mathbf{F}$ puede ser derivado como el gradiente de una función escalar (es decir, real; valuada en los reales) U. A U se le llama *función potencial.* Entonces

$$\mathbf{F} = \textbf{Grad}\ U = \nabla U \text{ en } \mathscr{E},$$

lo que quiere decir que $\dfrac{\partial U}{\partial x} = -A(x, y)$ y $\dfrac{\partial U}{\partial y} = -B(x, y)$ en $\mathscr{E}$. Así pues, decir que $\mathbf{F}$ es igual al gradiente de una potencial U es equivalente a afirmar que la diferencial $\Delta = \mathbf{F}(\mathbf{x}) \cdot d\mathbf{x}$ es exacta con $\mathbf{F}(\mathbf{x}) \cdot d\mathbf{x} = dU$. De donde, *si en $\mathscr{E}$ puede derivarse el campo* $\mathbf{F}$ *como el gradiente de una función potencial U, entonces el campo* $\mathbf{F}$ *es conservativo en $\mathscr{E}$ y el trabajo hecho al mover una partícula unitaria de* $\mathbf{P}$ *a* $\mathbf{Q}$ *a lo largo de una curva en $\mathscr{E}$ es igual a la diferencia en potencial entre los dos puntos.*

La introducción de la integral curvilínea y lo que conocemos acerca de ella arroja nueva luz sobre el teorema 12.11. Tal teorema afirma que, si A y B tienen primeras derivadas parciales continuas en un rectángulo abierto $\mathscr{R}$ entonces

13.6 $$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x} \quad \text{en } \mathscr{R}$$

es una condición suficiente para que $\Delta = A(x, y)dx + B(x, y)dy$ sea exacta sobre $\mathscr{R}$. Se demostró que si 13.6 se satisface en $\mathscr{R}$ entonces $\Delta = dF$ donde

$$F(x, y) = \int_{x_0}^{x} A(s, y_0)\, ds + \int_{y_0}^{y} B(x, s)\, ds.$$

Esto puede ahora interpretarse como una integral curvilínea a lo largo de una curva γ. Es la integral curvilínea de Δ a lo largo del segmento rectilíneo de (x_0, y_0) hasta (x, y_0) más el segmento rectilíneo de (x, y_0) hasta (x, y). La hipótesis de que $\mathscr{R}$ es un rectángulo abierto nos asegura que esta curva γ está en $\mathscr{R}$. Como Δ es exacta. la integral curvilínea que define F no depende

de la trayectoria y F puede calcularse a lo largo de cualquier curva lisa a trozos que se desee con tal de que esté en $\mathscr{R}$ y vaya del punto fijo (x_0, y_0) al (x, y).

13.7 Ejemplo. Resuélvase

$$\Delta = (x^2+4xy-y^2)dx+(2x^2-2xy+y^2)dy = 0$$

SOLUCIÓN.

$$\frac{\partial A}{\partial y} = 4x-2y = \frac{\partial B}{\partial x}$$

y por tanto Δ es exacta sobre R^2. Integrando a lo largo del segmento rectilíneo de $(0, 0)$ a (x, y):

$$\{(tx, ty) \mid t \in [0, 1]\},$$

entonces

$$\Delta = (t^2x^2+4txty-t^2y^2)xdt+(2t^2x^2-2txty+t^2y^2)ydt$$

y tenemos

$$\begin{aligned} F(x, y) &= \int_0^1 [t^2(x^2+4xy-y^2)x\,dt+t^2(2x^2-2xy+y^2)y\,dt] \\ &= \tfrac{1}{3}(x^2+4xy-y^2)x+\tfrac{1}{3}(2x^2-2xy+y^2)y \\ &= \tfrac{1}{3}x^3+2x^2y-xy^2+\tfrac{1}{3}y^3 . \end{aligned}$$

Por tanto $\frac{1}{3}x^3+2x^2y-xy^2+\frac{1}{3}y^3 = c$.

Probemos ahora que si 13.6 se verifica sobre un conjunto abierto convexo $\mathscr{E}$ entonces Δ es exacta sobre $\mathscr{E}$. (Un conjunto es *convexo* si el segmento rectilíneo que une un par cualquiera de puntos del conjunto está todo él contenido en el conjunto.)

13.8 Teorema. *Si A y B tienen derivadas parciales primeras continuas en un conjunto abierto y convexo $\mathscr{E}$ y si*

13.9
$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x} \quad en\ \mathscr{E},$$

entonces $\Delta = A(x, y)dx+B(x, y)dy$ es exacta en $\mathscr{E}$.

PRUEBA. Podemos suponer que el origen está en el conjunto $\mathscr{E}$. Si así no fuera podemos trasladar un punto de $\mathscr{E}$ al origen y esto no afecta la convexidad de $\mathscr{E}$ ni la exactitud de Δ. Definamos $F(x, y)$ como la integral

curvilínea de Δ a lo largo del segmento rectilíneo que va de $(0, 0)$ a (x, y). Entonces

$$F(x, y) = \int_0^1 [xA(tx, ty) + yB(tx, ty)]\,dt,$$

y

$$\frac{\partial F}{\partial x} = \int_0^1 [A(tx, ty) + txD_1A(tx, ty) + tyD_1B(tx, ty)]\,dt.$$

Usando 13.9 vemos que

$$\frac{\partial F}{\partial x} = \int_0^1 \frac{d}{dt}(tA(tx, ty))\,dt = A(x, y).$$

Análogamente $\quad \dfrac{\partial F}{\partial y} = B(x, y).$

13.10 Ejemplo. Pruébese que

$$\Delta = \sqrt{x^2 - y}\,[2y(4x^2 - y)dx + x(2x^2 - 5y)dy]$$

es exacta para $y \leqslant x^2$.

SOLUCIÓN. $\dfrac{\partial A}{\partial y} = \dfrac{\partial B}{\partial x} = (x^2 - y)^{-1/2}(8x^4 + 5y^2 - 16x^2y).$

La región $\mathscr{E}$ definida por $y < x^2$ está fuera de la parábola y no es un conjunto abierto convexo. En realidad no podemos encontrar ni un solo punto (x_0, y_0) de $\mathscr{E}$ con la propiedad de que el segmento rectilíneo de (x_0, y_0) a (x, y) esté en $\mathscr{E}$ para todo (x, y) en $\mathscr{E}$. Sin embargo, sabemos que alrededor de cada punto de $\mathscr{E}$ hay un círculo en el interior del cual 13.9 se satisface y, por tanto, Δ es localmente exacta (exacta en una vecindad de cada punto).

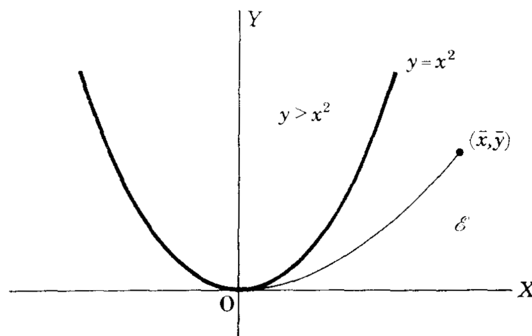


FIGURA 28

¿No será Δ, entonces, exacta en todo $\mathscr{E}$? (Esto no es en general cierto como veremos dentro de un momento.) Definamos $F(\bar{x}, \bar{y})$ como la integral curvilínea de Δ a lo largo de la parábola $x = \bar{x}t$, $y = \bar{y}t^2$ desde (0, 0) a $(\bar{x}, \bar{y})$ (figura 28).

$$F(\bar{x}, \bar{y}) = \sqrt{\bar{x}^2 - \bar{y}} \int_0^1 [(8\bar{y}\bar{x}^2 - 2\bar{y}^2)t^5 \bar{x}\,dt + (2\bar{x}^3 - 5\bar{x}\bar{y})2t^5 \bar{y}\,dt]$$

$$= 2\bar{x}\bar{y}(\bar{x}^2 - \bar{y})^{3/2}.$$

Es entonces fácil verificar que $dF = \Delta$ para $y \leqslant x^2$.

Hemos probado, pág. 671, que 13.9 es una condición necesaria para que Δ sea exacta sobre un conjunto abierto $\mathscr{E}$. Como señalamos en el anterior ejemplo, si esta condición se satisface en un conjunto abierto $\mathscr{E}$, entonces Δ es localmente exacta; es decir, Δ es exacta en una vecindad de cada punto. Plantea esto el problema de si 13.9 es una condición suficiente para que Δ sea exacta en $\mathscr{E}$ (exacta globalmente). El siguiente ejemplo muestra que en general esto no es cierto. Consideremos la forma diferencial

13.11 $$\Delta = \frac{x\,dy - y\,dx}{x^2 + y^2}, \qquad x^2 + y^2 \neq 0.$$

Se verifica fácilmente (problema 6) que 13.9 se satisface en todo el plano excluyendo el origen, y A y B tienen derivadas parciales continuas de todos los órdenes para $x^2 + y^2 \neq 0$. Calculando

$$\Delta(\gamma) = \int_\gamma \frac{x\,dy - y\,dx}{x^2 + y^2}$$

alrededor de la circunferencia γ: $x = a \cos \theta$, $y = a \operatorname{sen} \theta$, $\theta \in [0, 2\pi]$ obtenemos

$$\Delta(\gamma) = \int_0^{2\pi} (\cos^2 \theta + \operatorname{sen}^2 \theta)\,d\theta = \int_0^{2\pi} d\theta = 2\pi.$$

La integral curvilínea no se anula alrededor de ninguna circunferencia con centro en el origen y, por tanto, no puede ser exacta en ninguna región que contenga una circunferencia con centro en el origen.

Así pues, para resumir, sabemos que 13.9 *es una condición necesaria para la exactitud sobre conjuntos abiertos* $\mathscr{E}$, *pero no siempre es suficiente. Si* $\mathscr{E}$ *es un conjunto abierto convexo, entonces* 13.9 *es tanto necesaria como suficiente.* No proseguiremos con el estudio de este problema. El lector interesado puede consultar las referencias [6] y [16].

Problemas

1. Evalúese $\Delta(\gamma) = \int_\gamma [x^2\,dx + (x^2 - y^2)\,dy]$ a lo largo de

a) La recta γ de $(0, 0)$ a $(1, 1)$.
b) La parábola $x = y^2$ de $(0, 0)$ a $(1, 1)$.
c) El arco poligonal de $(0, 0)$ a $(0, 1)$ a $(1, 1)$.
d) El arco poligonal de $(0, 0)$ a $(1, 0)$ a $(1, 1)$.

2. Evalúese $\Delta(\gamma) = \int_\gamma [y^2\,dx + 2xy\,dy]$ a lo largo de las curvas del problema 1.

3. Evalúese $\Delta(\gamma) = \int_\gamma [xy\,dx + x^2\,dy]$ en dirección de las manecillas del reloj alrededor del cuadrado γ con vértices $(0, 0)$, $(0, 1)$, $(1, 1)$, $(1, 0)$.

4. Evalúese $\int_\gamma [xy\,dx - y^2\,dy]$ donde

a) $\gamma : x = t$, $y = t^2$, $t \in [0, 1]$.
b) γ es el segmento rectilíneo de $(0, 0)$ a $(1, 1)$.
c) γ es el triángulo de vértices sucesivos $(0, 0)$, $(0, 1)$, $(1, 0)$.

5. Evalúese $\int_\gamma (x\,dx + y\,dy)$ a lo largo de las curvas γ del problema 4.

6. Pruébese que la forma diferencial 13.11 satisface la condición 13.9 en todo el plano con excepción del origen.

7. Sean $\mathbf{P}_1 = r_1(\cos\theta_1, \operatorname{sen}\theta_1)$, $\mathbf{P}_2 = r_2(\cos\theta_2, \operatorname{sen}\theta_2)$ puntos en el semiplano superior; $\theta_1, \theta_2 \in \langle 0, \pi \rangle$, $r_1 > 0$, $r_2 > 0$. Sea γ una curva lisa a trozos situada en el semiplano superior que va de $\mathbf{P}_1$ a $\mathbf{P}_2$. Pruébese que

$$\int_\gamma \frac{x\,dy - y\,dx}{x^2 + y^2} = \theta_2 - \theta_1 .$$

8. Evalúese $\int_\gamma \frac{x\,dy - y\,dx}{x^2 + y^2}$ donde γ es cualquier curva simple cerrada del plano que no pasa por el origen.

9. Pruébese que el campo de fuerza $\mathbf{F}(x, y) = (\operatorname{sen} y, x\cos y)$ es conservativo y calcúlese el trabajo hecho al mover una partícula del punto $(0, \pi)$ al punto $(2\pi, -\pi)$.

10. Determínese cuál o cuáles de los siguientes campos de fuerza son

conservativos. Cuando sea posible exprésese el campo de fuerza como el gradiente de una función potencial.

a) $\mathbf{F}(x, y) = (3x^2y^2 - 3y^2, 2x^3y - 6xy)$
b) $\mathbf{F}(x, y) = (y^2 \cos xy + x^3, xy \cos xy + \operatorname{sen} xy)$
c) $\mathbf{F}(x, y) = (2x \operatorname{sen}(x-y), x^2 \operatorname{sen}(x-y))$
d) $\mathbf{F}(x, y) = e^x(2x + x^2 + y^2, 2y)$.

11. Sean

$$\gamma: \; x = u(t), \quad y = v(t), \quad t \in [a, b] = \mathscr{I}$$
$$\gamma^*: \; x = u^*(t), \quad y = v^*(t), \quad t \in [a^*, b^*] = \mathscr{I}^*$$

un par de curvas lisas. Si hay una función φ con las siguientes propiedades: 1) φ transforma $\mathscr{I}$ sobre $\mathscr{I}^*$ ($\varphi(\mathscr{I}) = \mathscr{I}^*$), 2) φ tiene una derivada continua positiva sobre $\mathscr{I}$, y 3) $(u^*(\varphi(t)), v^*(\varphi(t))) = (u(t), v(t))$ para todo t en $\mathscr{I}$, entonces las curvas γ y γ^* se dice que son *diferencialmente equivalentes*. Pruébese que: si γ y γ^* son diferencialmente equivalentes, entonces $\Delta(\gamma) = \Delta(\gamma^*)$.

12. Supongamos que A y B tienen primeras derivadas parciales continuas sobre un conjunto convexo y abierto $\mathscr{E}$.

a) Pruébese que $\mu(x, y)$ es un factor de integración de

$$\Delta = A(x, y)dx + B(x, y)dy$$

si y sólo si μ no se anula y es una solución sobre $\mathscr{E}$ de

$$\mu\left(\frac{\partial A}{\partial y} - \frac{\partial B}{\partial x}\right) = B\frac{\partial \mu}{\partial x} - A\frac{\partial \mu}{\partial y}.$$

b) Obténgase una condición necesaria y suficiente para que $\mu(x)$ sea un factor de integración de Δ. ¿Qué es el factor de integración?

c) A se dice que es *homogénea de grado k sobre* $\mathscr{E}$ si $A(tx, ty) = t^k A(x,y)$ para todo (x, y) en $\mathscr{E}$ y todo t con la propiedad de que (tx, ty) esté en $\mathscr{E}$. Si A y B son, ambas, homogéneas de grado k sobre $\mathscr{E}$, entonces se dice que Δ es homogénea de grado k. Bajo las hipótesis de que Δ es homogénea de grado k sobre $\mathscr{E}$ y de que

$$\mu(x, y) = xA(x, y) + yB(x, y)$$

no se anula en $\mathscr{E}$, pruébese que μ^{-1} es un factor de integración (sobre $\mathscr{E}$) de Δ.

d) Si Δ es tanto exacta como homogénea sobre $\mathscr{E}$, resuélvase $\Delta = 0$.

14. CURVAS INTEGRALES

En esta sección queremos discutir un sistema

$$\dot{x} = P(x, y)$$
$$\dot{y} = Q(x, y)$$

de dos ecuaciones diferenciales de primer orden para las funciones desconocidas x y y. Para relacionarlo con lo que antes hemos estado viendo reemplazamos P por B y Q por $-A$. El sistema es, entonces,

14.1
$$\dot{x} = B(x, y)$$
$$\dot{y} = -A(x, y),$$

y suponemos en todo lo que sigue que A y B son continuas en un conjunto abierto $\mathscr{E}$ de R^2. El espacio R^2 se llama *espacio fase* o *espacio estado*. Cada punto (x, y) del espacio fase se corresponde con un *estado* del sistema. Las ecuaciones diferenciales 14.1 describen la manera en que el sistema cambia con el tiempo. Una solución $x = u(t)$, $y = v(t)$, $t \in \langle a, b \rangle$, define una curva γ en el espacio fase.

Piénsese del sistema 14.1 como si definiera un flujo en el espacio fase. En cada punto (x, y) la velocidad de una partícula en el flujo es $(\dot{x}, \dot{y}) = (B(x, y), -A(x, y))$. Una solución describe el movimiento de una partícula en el flujo y la curva γ es una trayectoria de una partícula en el flujo. Toda solución de 14.1 es obviamente una solución de

14.2
$$A(x, y)\dot{x} + B(x, y)\dot{y} = 0,$$

pero lo recíproco no es cierto. No toda solución de 14.2 es una solución de 14.1. La razón de esto es que 14.2 solamente afirma que la curva γ definida por una solución de 14.2 está en cada punto (x, y) en la dirección $(B(x, y), -A(x, y))$. Nada se dice acerca de la magnitud de la velocidad del movimiento en esa dirección. El campo vectorial $(B(x, y), -A(x, y))$ es un "campo direccional" según 14.2, pero un "campo de velocidades" según 14.1.

Para sistemas no lineales no siempre podemos ser capaces de dar soluciones generales en términos de funciones elementales, ni aun en términos de funciones de aquéllas de cualquier clase para las que se hayan calculado tablas o cuyas propiedades hayan sido estudiadas con anterioridad. Es, sin embargo, posible a menudo obtener una gran cantidad de información acerca de las soluciones sin resolver las ecuaciones. Lo que estamos a punto de discutir e ilustrar es cómo puede obtenerse información acerca de las trayectorias definidas por las soluciones incluso cuando estas últimas no nos son conocidas.

14.3 Definición. *Una función real $F(x, y)$ definida sobre un conjunto abierto $\mathscr{E}$ de* R^2 *se dice que es una* ***integral*** *de* 14.1 *si F es constante a lo largo de toda solución. Esto significa que si $(u(t), v(t))$ es una solución de* 14.1 *en $\mathscr{E}$*

para $t \in \langle a, b \rangle$ *entonces* $F(u(t), v(t)) = c$ *para alguna constante* c *y para todo* $t \in \langle a, b \rangle$. *Las curvas de nivel* $F(x, y) = c$ *se llaman* ***curvas integrales*** *de* 14.1.

Cualquier función F que es constante sobre $\mathscr{E}$ es, desde luego, una integral, pero es una integral trivial sin ningún interés. Probaremos ahora que una solución $F(x, y) = c$ de $A(x, y)dx + B(x, y)dy = 0$ nos da una integral de 14.1.

14.4 Teorema. *Supongamos que* A *y* B *son continuas sobre un conjunto abierto* $\mathscr{E}$ *de* R^2. *Si* $F(x, y) = c$ *es una solución en* $\mathscr{E}$ *de*

14.5
$$A(x, y)dx + B(x, y)dy = 0,$$

entonces F *es una integral sobre* $\mathscr{E}$ *de*

$$\begin{aligned} \dot{x} &= B(x, y) \\ \dot{y} &= -A(x, y). \end{aligned}$$

PRUEBA. Como toda solución de 14.1 es también una solución de 14.2, este teorema es una consecuencia de lo que probamos en la sección anterior. Una prueba directa es también simple e instructiva. Nuestras hipótesis sobre F implican que

$$\dot{x} = B(x, y) = \mu^{-1}(x, y)\,\frac{\partial F(x, y)}{\partial y}$$

$$\dot{y} = -A(x, y) = -\mu^{-1}(x, y)\,\frac{\partial F(x, y)}{\partial x}$$

donde μ no se anula en $\mathscr{E}$. Por tanto, a lo largo de las soluciones de 14.1

$$\frac{dF}{dt} = \frac{\partial F}{\partial x}\dot{x} + \frac{\partial F}{\partial y}\dot{y} = 0.$$

F es constante a lo largo de las soluciones y es una integral. Y esto completa la prueba.

Estudiemos el movimiento de una partícula a lo largo de una recta y sujeta a la fuerza restauradora $x + x^2$, donde x es la distancia dirigida de la partícula a un punto 0. La segunda ley de Newton nos dice que, entonces,

$$m\ddot{x} + x + x^2 = 0.$$

Con $y = m\dot{x}$, el momento, esta ecuación es equivalente al sistema

14.6
$$\begin{aligned} m\dot{x} &= y \\ \dot{y} &= -x - x^2. \end{aligned}$$

La ecuación diferencial para una integral es

$$(x+x^2)dx+m^{-1}ydy = 0.$$

De donde vemos que $H(x, y) = \frac{1}{2}x^2 + \frac{1}{3}x^3 + \frac{1}{2m}y^2$ es una integral y las soluciones de 14.6 se encuentran sobre las curvas integrales $H(x, y) = c$. Las curvas integrales se muestran en la figura 29. Físicamente $H(x, y)$ es la energía total del sistema. La cantidad $\frac{1}{2}x^2+\frac{1}{3}x^3$ es la energía potencial del sistema y $\frac{1}{2m}y^2 = \frac{1}{2}m\dot{x}^2$ es la energía cinética. A H se le llama *integral de energía*, y el hecho de que H sea constante es la ley de conservación de la energía. Por la figura 29 vemos que un conocimiento de las curvas integrales nos da una gran cantidad de información acerca del comportamiento del sistema. Las flechas muestran la dirección de las soluciones cuando el tiempo transcurre. Esto se determina fácilmente por 14.6. Por encima del eje X, $\dot{x} > 0$ y x es creciente. Por debajo del eje X, $\dot{x} < 0$ y x es decreciente. Nótese particularmente la diferencia entre el comportamiento del sistema cerca de $(0, 0)$ y de $(-1, 0)$. Estos estados son "estados de equilibrio" del sistema: $(0, 0)$ y $(-1, 0)$ son soluciones de 14.6. En estos puntos el campo de velocidades $(y, -x-x^2)$ se anula. Si el sistema se inicia en cualquiera de estos dos estados, permanece en ese estado. Si el sistema está en el estado $(0, 0)$ y se perturba ligeramente, se moverá a una de las curvas integrales vecinas, pero permanecerá cerca del estado de equilibrio. El estado de equilibrio es "estable" y por las curvas cerradas a su alrededor se llama un "centro". Las curvas cerradas corresponden a soluciones periódicas (problema 6, pág. 707). Si la energía total no es

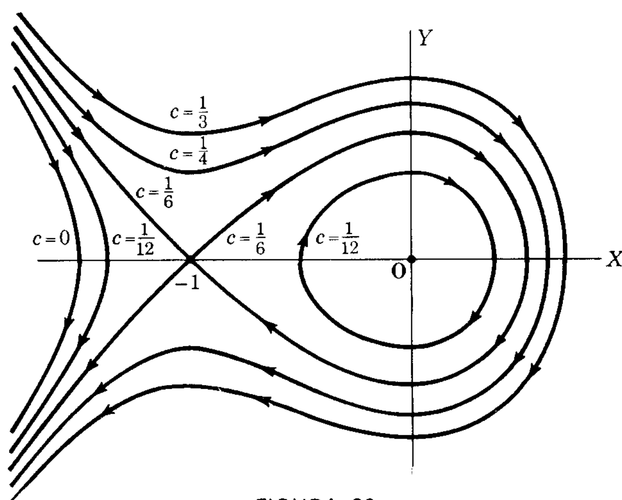


FIGURA 29

demasiado grande, el sistema oscila alrededor del estado de equilibrio. Para valores grandes de la energía total las soluciones pueden escapar y tender al infinito cuando t tiende al infinito (problema 3, pág. 693). Si el sistema está en el estado de equilibrio $(-1, 0)$ y es perturbado ligeramente, no necesariamente permanecerá cerca del estado de equilibrio. Con la excepción de las dos curvas integrales que tienden al estado de equilibrio, una perturbación será causa de que el sistema escape u oscile alrededor de $(0, 0)$. A causa de la naturaleza de las curvas integrales cerca de $(-1, 0)$, este estado de equilibrio se llama "punto de ensilladura", y es "inestable". Los valores máximos de las oscilaciones de $x(t)$ corresponden en donde las curvas integrales cerradas intersectan el eje X. Para determinar los periodos de estas oscilaciones se necesitaría un análisis más profundo.

Este ejemplo y el que a continuación daremos nos muestran la riqueza de información que puede obtenerse sin resolver la ecuación, y sirve también para demostrar la complejidad de comportamiento que puede esperarse en los sistemas no lineales. Ambas ecuaciones (la 14.6 y la ecuación del ejemplo 14.7) tienen la misma aproximación lineal $\dot{x} = y$, $\dot{y} = -x$, pero se comportan muy diferentemente. La aproximación lineal resulta que nos da en ambos casos información acerca del comportamiento cerca del origen pero ni incluso esto sucede en general. Por ejemplo, $\dot{x} = y$, $\dot{y} = -x + y^3$ tiene un estado de equilibrio único en el origen, que es inestable, y toda solución tiende a infinito cuando t tiende a infinito. Hay casos en que una aproximación lineal da alguna información local acerca del comportamiento de los sistemas no lineales. Lo que aquí hemos dicho debe, sin embargo, hacernos ver el cuidado que es necesario tener al usar aproximaciones lineales.

14.7 Ejemplo. Obténgase una integral de

$$\begin{aligned} \dot{x} &= y(1-x^2) \\ \dot{y} &= -x(1-y^2) \end{aligned}$$

y dibújense las curvas integrales.

SOLUCIÓN. Consideremos la ecuación diferencial

$$x(1-y^2)dx + y(1-x^2)dy = 0.$$

La ecuación es separable. Separando variables obtenemos

$$\frac{x\,dx}{1-x^2} + \frac{y\,dy}{1-y^2} = 0,$$

$$\ln|1-x^2| + \ln|1-y^2| = \ln c,$$

$$(1-x^2)(1-y^2) = c.$$

En lugar de dejarnos arrastrar por el hecho de que la ecuación es separable

habría sido mejor que hubiésemos notado inmediatamente que la ecuación es exacta. Obtenemos entonces la integral $(1-x^2)(1-y^2)$ sin que se nos presenten problemas acerca de la anulación de $(1-x^2)(1-y^2)$. Las curvas integrales aparecen en la figura 30. El sistema tiene cinco estados de equilibrio obtenidos al igualar el campo de velocidades $(y(1-x^2), -x(1-y^2))$ a cero. Son: (0, 0), (1, 1), (−1, 1), (−1, −1) y (1, −1). Otro término para designar los estados de equilibrio es "punto crítico". Como en el estudio de los valores máximo y mínimo de las funciones, hay también puntos críticos

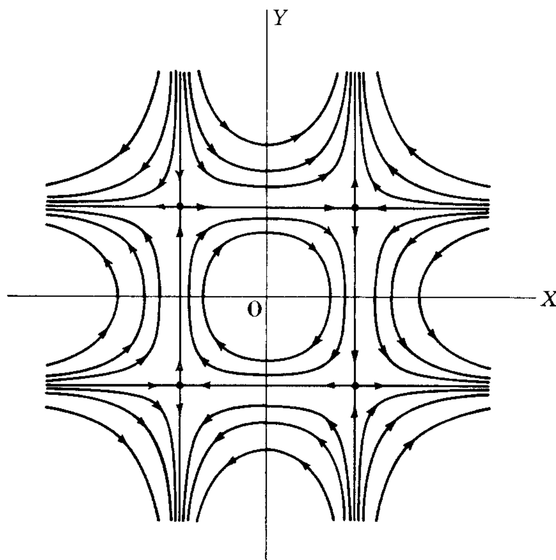


FIGURA 30

en la integral. Las curvas cerradas alrededor del origen corresponden a soluciones periódicas. Si el estado inicial está en el interior del cuadrado, el sistema oscila alrededor del origen. Fuera del cuadrado las soluciones, con excepción de aquellas que tienden al estado de equilibrio, tienden a infinito cuando t tiende a infinito. Estas soluciones excepcionales son altamente inestables. Una ligera perturbación lleva el sistema a una solución que tiende al infinito u oscila. Los cuatro estados de equilibrio en los vértices del cuadrado son puntos de ensilladura y son inestables.

En el próximo ejemplo estudiaremos el clásico problema de mecánica del movimiento de un péndulo. Veremos cómo un conocimiento de las curvas integrales nos da alguna información cuantitativa y una buena visión de los aspectos cualitativos del comportamiento de un péndulo.

14.8 Ejemplo. Discútase el movimiento de un péndulo simple despreciando la fricción.

SOLUCIÓN. Consideremos un péndulo de longitud l y masa m (figura 31).

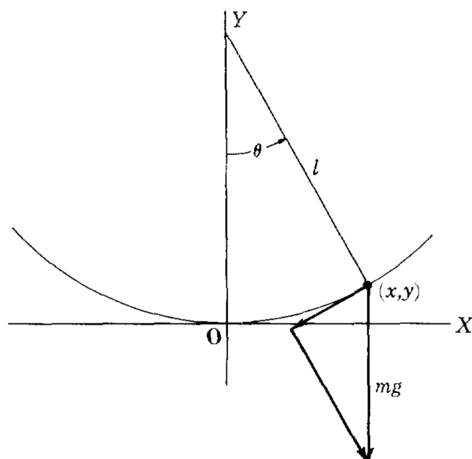


FIGURA 31

Sea $((x(t), y(t))$ la posición de la partícula en el instante t. De acuerdo con la segunda ley del movimiento de Newton

$$m\mathbf{a} = m(\ddot{x}, \ddot{y}) = (0, -mg)$$

sujeto a la constricción $x^2+y^2 = l^2$. Ahora bien,

$$x = l \operatorname{sen} \theta$$
$$y = l - l \cos \theta,$$

de modo que

$$\ddot{x} = -l\dot{\theta}^2 \operatorname{sen} \theta + l\ddot{\theta} \cos \theta,$$
$$\ddot{y} = l\dot{\theta}^2 \cos \theta + l\ddot{\theta} \operatorname{sen} \theta.$$

De donde

$$l\ddot{\theta} = \cos \theta \, \ddot{x} + \operatorname{sen} \theta \, \ddot{y} = -g \operatorname{sen} \theta$$

y la ecuación diferencial para el ángulo θ es

$$l\ddot{\theta} + g \operatorname{sen} \theta = 0.$$

Haciendo el cambio $t = \alpha\tau$ en la unidad de tiempo, y $u(\tau) = \theta(\alpha\tau)$, la ecuación diferencial para el ángulo como función del nuevo tiempo es

$$\alpha^{-2} l u'' + g \operatorname{sen} u = 0.$$

Con $\alpha^2 = l/g$ la ecuación diferencial para el ángulo u es

$$u'' + \operatorname{sen} u = 0.$$

El ángulo u y el tiempo τ son cantidades sin dimensión. Podemos ver tanto

directamente en la ecuación diferencial como en el sistema equivalente

$$u' = v$$
$$v' = -\operatorname{sen} u,$$

que $vv' + u' \operatorname{sen} u = D_\tau(\frac{1}{2}v^2 + 1 - \cos u) = 0$ y

$$\tfrac{1}{2}v^2 + (1 - \cos u)$$

es una integral y de nuevo una integral de energía. Las curvas integrales pueden escribirse

$$v^2 + 4 \operatorname{sen}^2 \frac{u}{2} = 2c^2 .$$

Estas curvas integrales están dibujadas en la figura 32. Observemos ahora

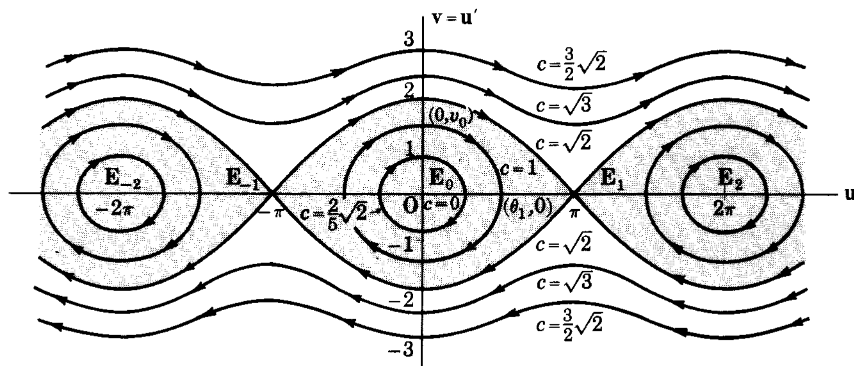


FIGURA 32

lo siguiente:

1) Los estados de equilibrio, que corresponden a $v = 0$, $\operatorname{sen} u = 0$, están en los puntos $\mathbf{E}_k = (k\pi, 0)$; k, un entero cualquiera. Físicamente éstos corresponden a los dos estados de equilibrio del sistema. Los puntos $\mathbf{E}_{2k}$ corresponden a las posiciones más bajas del péndulo con el péndulo inmóvil. Estos son centros y son estables. Los puntos $\mathbf{E}_{2k+1}$ corresponden al péndulo inmóvil en su posición más alta ($\theta = \pi$). Estos puntos son puntos de ensilladura y son inestables. Esto confirma lo que sabemos físicamente acerca de un péndulo en equilibrio en la posición vertical $\theta = \pi$.

2) Si el estado inicial de los sistemas se encuentra cerca del origen dentro del área sombreada de la figura 32 ($c^2 < 2$) el péndulo oscila para atrás y adelante alrededor de su estado de equilibrio $\mathbf{E}_0$. En el caso llamado de las "pequeñas oscilaciones", donde u permanece pequeño, podemos aproximar $\operatorname{sen} u$ por u y las curvas integrales son casi circunferencias:

$v^2+u^2=2c^2$. La aproximación lineal de la ecuación diferencial para pequeñas oscilaciones es $u''+u=0$, y

$$u(\tau) = a \operatorname{sen}(\tau+\delta), \quad v(\tau) = a\cos(\tau+\delta), \quad a = \sqrt{2c}.$$

El periodo de las pequeñas oscilaciones es aproximadamente 2π. En tiempo real t el periodo es $2\pi\sqrt{l/g}$.

3) A medida que el estado inicial del péndulo se aproxima a la frontera de la región sombreada (que c^2 tiende a 2), las curvas integrales tienden a tomar la forma de la frontera $\left(v^2 = 4\cos^2\frac{u}{2}\right)$, y no pueden seguir aproximándose por circunferencias.

4) La frontera de la región sombreada corresponde a las condiciones iniciales para las que el péndulo se aproxima al estado de equilibrio $\mathbf{E}_1$ o $\mathbf{E}_{-1}$. Podemos esperar a medida que las oscilaciones se aproximan a la frontera que los periodos de las oscilaciones tiendan a infinito. Más adelante, en 7), hablaremos más de esto.

5) Si el estado inicial se encuentra fuera de la región sombreada, entonces el péndulo deja de oscilar adelante y atrás y entonces gira. Las revoluciones son en dirección contraria a las de las manecillas del reloj por encima de la región sombreada y en la dirección de las manecillas del reloj debajo.

6) Los valores extremos de u (las amplitudes de las oscilaciones) ocurren donde las curvas integrales cortan al eje X ($v=u'=0$), y los valores extremos de u' pueden determinarse por las intersecciones de las curvas integrales con el eje v ($u=0$) y la recta vertical $u=\pi$.

7) Supongamos que $u(0)=0$, $v(0)=u'(0)=v_0$; donde $0<v_0<2$, de manera que el péndulo oscila. Sea $(\theta_1, 0)$ el punto donde la curva integral en la dirección de τ creciente encuentra por primera vez el eje u. Entonces θ_1 es la amplitud de la oscilación. Sea τ_1 el tiempo para que el sistema vaya del estado $(0, v_0)$ al estado $(\theta, 0)$. Entonces $u(\tau_1)=\theta_1$ y por razones de simetría el periodo de la oscilación es $4\tau_1$, o en tiempo real, el periodo es $T = 4\sqrt{\frac{l}{g}}\,\tau_1$. Queremos ahora obtener una fórmula para el periodo como una función de la amplitud θ. Por la integral de energía tenemos

$$v^2(\tau)+4\operatorname{sen}^2\frac{u(\tau)}{2} = 2c^2 = v_0{}^2 = 4\operatorname{sen}^2\frac{u(\tau_1)}{2} = 4\operatorname{sen}^2\frac{\theta_1}{2},$$

donde $u(\tau)$, $v(\tau)$ son las soluciones correspondientes a $u(0)=0$, $v(0)=v_0$. De donde

$$u'(\tau) = v(\tau) = \left(v_0{}^2 - 4\operatorname{sen}^2\frac{u(\tau)}{2}\right)^{1/2},$$

$$= 2\left(k^2 - \operatorname{sen}^2\frac{u(\tau)}{2}\right)^{1/2},$$

donde $k = \frac{1}{2}v_0 = \operatorname{sen} \frac{1}{2}\theta_1$. De esto obtenemos

$$\frac{1}{2}\int_0^{u(\tau)} \frac{d\alpha}{(k^2 - \operatorname{sen}^2 \frac{1}{2}\alpha)^{1/2}} = \tau.$$

Haciendo el cambio de variable $\operatorname{sen} \dfrac{\alpha}{2} = k \operatorname{sen} \varphi$, vemos que

$$\tau = \int_0^{h(\tau)} \frac{d\varphi}{(1 - k^2 \operatorname{sen}^2 \varphi)^{1/2}},$$

donde $k \operatorname{sen} h(\tau) = \operatorname{sen} \dfrac{u(\tau)}{2}$. Esta integral es una integral elíptica de primera clase, y la función definida por la integral es

$$F(z, k) = \int_0^{z} \frac{d\varphi}{(1 - k^2 \operatorname{sen}^2 \varphi)^{1/2}}.$$

Existen tablas muy extensas de esta función y de acuerdo con ellas podemos calcular los tiempos empleados en moverse de un punto a otro a lo largo de las curvas integrales. Recíprocamente, dada τ podemos calcular el ángulo $u(\tau)$. En particular, como $u(\tau_1) = \theta_1$, $\operatorname{sen} h(\tau_1) = 1$ y $h(\tau_1) = \dfrac{\pi}{2}$. El

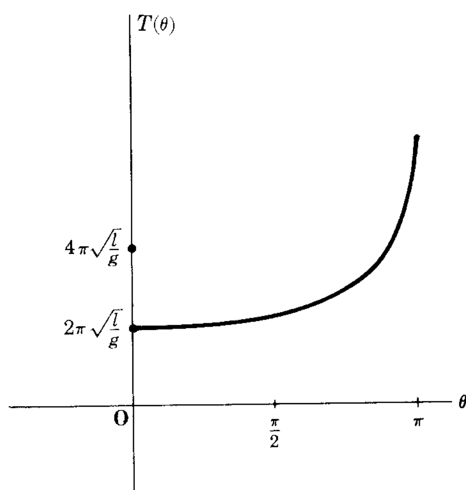


FIGURA 33 Periodo ante amplitud

periodo de la oscilación de amplitud θ_1, $(0 < \theta_1 < \pi)$ es $4\tau_1 = 4F\left(\frac{\pi}{2}, \operatorname{sen}\frac{\theta_1}{2}\right)$. En tiempo real t el periodo $T(\theta_1)$ es

$$T(\theta_1) = 4\sqrt{\frac{l}{g}}\,F\left(\frac{\pi}{2}, \operatorname{sen}\frac{\theta_1}{2}\right).$$

Se sigue del problema 10, página 567, que $T(\theta)$ tiende a ∞ cuando θ_1 tiende a π. Como $F\left(\frac{\pi}{2}, 0\right) = \frac{\pi}{2}$, y F es continua, $2\pi\sqrt{\frac{l}{g}}$ es una buena aproximación para pequeñas oscilaciones (figura 33). $F\left(\frac{\pi}{2}, k\right)$ se llama integral elíptica *completa* de primera clase.

Problemas

1. Escríbase cada una de las siguientes como un sistema de dos ecuaciones de primer orden, obténgase una integral y dibújense las curvas integrales:

a) $\ddot{x} + x = 0$ *b*) $\ddot{x} - x = 0$

c) $\ddot{x} + x + x^3 = 0$ *d*) $\ddot{x} - x + x^3 = 0$

e) $\ddot{x} + \operatorname{sen} = 0$ *f*) $\dot{x} = x(2y^3 - x^3)$
$\dot{y} = -y(2x^3 - y^3)$

2. La ecuación de un oscilador (ecuación de Van der Pol) es

$$\ddot{x} + \mu(1 - x^2)\dot{x} + x = 0$$

Pruébese que

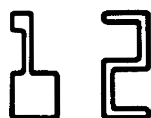
$$\dot{x} = \mu[y - (x - \tfrac{1}{3}x^3)]$$

$$\dot{y} = -\frac{1}{\mu}x$$

es un sistema equivalente. Discútanse las trayectorias definidas por soluciones en el caso en que μ es muy grande por un estudio del flujo definido por la ecuación diferencial en el plano fase. Una gran μ resulta en las que se llaman "oscilaciones de relajación".

3. Demuéstrese que hay soluciones de 14.6 que tienden a ∞ cuando $t \to \infty$.





12 Funciones definidas por ecuaciones diferenciales

1. INTRODUCCIÓN

Continuando con nuestro estudio introductorio de ecuaciones diferenciales, abordaremos en primer lugar el problema de la existencia y unicidad de soluciones de las ecuaciones diferenciales. Al hacer esto introduciremos dos tópicos de considerable importancia matemática: los teoremas de punto fijo y las aproximaciones sucesivas. Probaremos primero un teorema de punto fijo para espacios vectoriales finitodimensionales y luego mostraremos cómo es que este teorema puede extenderse a espacios en que los elementos son funciones. Estos espacios de funciones son espacios vectoriales infinitodimensionales, pero vemos que para estos problemas la dimensión del espacio no juega papel alguno. El teorema del punto fijo para espacios de funciones nos da un teorema de existencia y unicidad

para las soluciones de las ecuaciones diferenciales. Ilustraremos después cómo es que este teorema de existencia y unicidad nos lleva a la definición de funciones por ecuaciones diferenciales, y aprenderemos un poco acerea de cómo podemos estudiar las propiedades de funciones definidas por ecuaciones diferenciales. El capítulo se concluye con un estudio de un sujeto íntimamente relacionado con los tratados: las series de Fourier y las aproximaciones de Fourier. Aquí, de nuevo veremos una analogía entre espacios de funciones y espacios vectoriales finitodimensionales.

2. TEOREMA DE PUNTO FIJO: APROXIMACIONES SUCESIVAS

El método de Newton para resolver la ecuación

$$f(x) = x^2 - 2 = 0$$

consiste en convertir el problema en él al equivalente por resolver

$$x = T(x) = x - \frac{f(x)}{f'(x)} = \frac{x}{2} + \frac{1}{x}, \quad \text{donde } f(x) = x^2 - 2.$$

Una solución de $x = T(x)$ se llama *punto fijo* de T; comenzaremos con una aproximación inicial x_0 y definiremos una sucesión de lo que esperamos sean aproximaciones sucesivas de un punto fijo por

$$\begin{aligned} x_1 &= T(x_0) \\ x_2 &= T(x_1) = T^{[2]}(x_0) \\ &\vdots \\ x_n &= T(x_{n-1}) = T^{[n]}(x_0). \end{aligned}$$

Por ejemplo, con $T(x) = \frac{x}{2} + \frac{1}{x}$ y $x_0 = 1$, obtenemos

$$\begin{aligned} x_1 &= T(1) = 1.5 \\ x_2 &= T(1.5) = 1.417 \\ x_3 &= T(1.42) = 1.41425 \end{aligned}$$

y tenemos ya una buena aproximación (con seis cifras significativas $\sqrt{2} = 1.41421$). ¿Por qué y cuándo trabaja esto? Notemos primero que si T es continuo sobre un intervalo cerrado $\mathscr{I}$ y si la sucesión de aproximaciones sucesivas está en $\mathscr{I}$ y converge, entonces el límite de la sucesión es un punto fijo de T. Sea $x_n \to \bar{x}$ cuando $n \to \infty$. Entonces

$$\bar{x} = \lim_{n\to\infty} x_n = \lim_{n\to\infty} T(x_{n-1}) = T(\lim_{n\to\infty} x_{n-1}) = T(\bar{x}).$$

Obtendremos ahora un teorema de punto fijo excepcionalmente simple, pero muy importante y sorprendentemente poderoso.

2.1 Teorema. *Una función real T definida sobre un intervalo $\mathscr{I}$ de* R *se dice que es una* ***contracción*** *sobre $\mathscr{I}$ si*

$$|T(x)-T(y)| \leqslant \lambda |x-y|$$

para todo x, y en $\mathscr{I}$ y algún $\lambda \in \langle 0, 1 \rangle$.

Una contracción transforma todo par de puntos x, y en un par $T(x)$, $T(y)$ que son más próximos. Ciertamente, si T es una contracción sobre $\mathscr{I}$ entonces T es continua sobre $\mathscr{I}$ (problema 1). El recíproco no es cierto. Sin embargo, *si T es diferenciable sobre $\mathscr{I}$ y si $|T'(x)| \leqslant \lambda < 1$ sobre $\mathscr{I}$, entonces T es una contracción sobre $\mathscr{I}$* (problema 2).

2.2 Teorema. *Si T es una contracción sobre un intervalo cerrado $\mathscr{I}$ y si T transforma $\mathscr{I}$ en sí mismo ($T(\mathscr{I}) \subset \mathscr{I}$), entonces T tiene un único punto fijo $\bar{x}$ en $\mathscr{I}$. Más específicamente, si x_0 es un punto cualquiera en $\mathscr{I}$, entonces $x_n = T^{[n]}(x_0) \to \bar{x}$ cuando $n \to \infty$ y*

2.3
$$|\bar{x}-x_n| \leqslant \frac{\lambda}{1-\lambda} |x_n - x_{n-1}|.$$

Prueba. Como la transformación se contrae sobre $\mathscr{I}$, no puede tener dos puntos fijos en $\mathscr{I}$. Supongamos que $\bar{x}_1$ y $\bar{x}_2$ son puntos fijos en $\mathscr{I}$. Entonces

$$|\bar{x}_1 - \bar{x}_2| = |T(\bar{x}_1) - T(\bar{x}_2)| \leqslant \lambda |\bar{x}_1 - \bar{x}_2|.$$

Como $\lambda < 1$, $\bar{x}_1 = \bar{x}_2$. Necesitamos ahora probar la existencia de un punto fijo. Para hacer esto tomamos en $\mathscr{I}$ un punto x_0 cualquiera y definimos la sucesión de aproximaciones sucesivas $x_n = T(x_{n-1}) = T^{[n]}(x_0)$. Entonces

$$|x_n - x_{n-1}| = |T(x_{n-1}) - T(x_{n-2})| \leqslant \lambda |x_{n-1} - x_{n-2}|.$$

Como $\lambda < 1$, la sucesión converge (problema 2, pág. 480). Sea $\lim_{n\to\infty} x_n = \bar{x}$, como $\mathscr{I}$ es cerrado, $\bar{x}$ está en $\mathscr{I}$. Entonces, como T es continua sobre $\mathscr{I}$,

$$\bar{x} = \lim_{n\to\infty} x_n = \lim_{n\to\infty} T(x_{n-1}) = T(\bar{x}).$$

Por tanto $\bar{x}$ es el punto fijo en $\mathscr{I}$. Todo lo que queda por hacer es establecer 2.3. Tenemos

$$|\bar{x} - x_n| = |T(\bar{x}) - T(x_{n-1})| \leqslant \lambda |\bar{x} - x_{n-1}| \leqslant \lambda [|x_n - x_{n-1}| + |\bar{x} - x_n|].$$

Por tanto

$$(1-\lambda)\, |\bar{x} - x_n| \leqslant \lambda |x_n - x_{n-1}|$$

y 2.3 se sigue.

La desigualdad 2.3 nos da una estimación útil del error en la n-ésima aproximación en términos de la diferencia entre esta aproximación y la precedente y nos dice cuándo podemos parar el cálculo. (Veáse también el problema 4.)

Nuestro propósito aquí no es proseguir con el problema de resolver ecuaciones $f(x) = 0$. Nuestra finalidad principal es introducir en la forma más sencilla posible el método de las aproximaciones sucesivas. Muchos problemas del análisis pueden reducirse a problemas de punto fijo de transformaciones de funciones en lugar de números. Entre éstos está el problema de resolver una ecuación diferencial. Queremos ahora extender el teorema 2.2 y el método de aproximaciones sucesivas a las transformaciones de funciones.

Sea $\mathscr{I}$ el intervalo cerrado $\mathscr{I} = [t_0 - a, t_0 + a]$ y denotemos por $\mathscr{C}$ el conjunto de todas las funciones continuas $\mathbf{x}$ de $\mathscr{I}$ en R^n. Para cada $\mathbf{x}$ en $\mathscr{C}$ definimos

2.4 $$\|\mathbf{x}\| = \text{Máx}\,\{|\mathbf{x}(t)|\ | t \in \mathscr{I}\}.$$

Esta función real sobre $\mathscr{C}$ se llama *norma*, y $\|\mathbf{x}\|$ se lee "norma de $\mathbf{x}$". Es lo correspondiente a un valor absoluto y es una medida de la distancia de una función al origen. La distancia entre dos funciones $\mathbf{x}$ y $\mathbf{y}$ en $\mathscr{C}$ es $\mathbf{x} - \mathbf{y}$. La norma tiene las propiedades características de un valor absoluto (problema 10):

2.5 $$\|\mathbf{x}\| \geqslant 0; \ \|\mathbf{x}\| = 0 \quad \text{si y sólo si } \mathbf{x} = \mathbf{0}.$$

2.6 $$\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$$

2.7 $$\|c\mathbf{x}\| = |c|\ \|\mathbf{x}\|.$$

Para una sucesión de funciones $\{\mathbf{x}^n\}$ en $\mathscr{C}$ definimos $\lim\limits_{n \to \infty} \mathbf{x}^n = \mathbf{x}$ si

$$\lim_{n \to \infty} \|\mathbf{x}^n - \mathbf{x}\| = 0.$$

Así pues, *la convergencia de una sucesión de funciones en $\mathscr{C}$ es la convergencia uniforme de la sucesión sobre el intervalo $\mathscr{I}$.*

Necesitamos ahora considerar transformaciones T de $\mathscr{C}$ en sí mismo. Ejemplos de tales transformaciones son $\mathbf{y} = T(\mathbf{x})$ donde

2.8 $$\mathbf{y} = A\mathbf{x}$$

2.9 $$\mathbf{y}(t) = \mathbf{F}(\mathbf{x}(t)) \quad (T\mathbf{x} = \mathbf{F} \circ \mathbf{x}),$$

2.10 $$\mathbf{y}(t) = \int_0^1 \varphi(s, t)\, \mathbf{f}(\mathbf{x}(s), s)\, ds,$$

2.11 $$\mathbf{y}(t) = \mathbf{x}^0 + \int_{t_0}^{t} \mathbf{f}(\mathbf{x}(\tau), \tau)\, d\tau.$$

La última transformación es la de mayor interés para nosotros por el momento. Si $\mathbf{x}$ es un punto fijo de la transformación T definida por 2.11 ($T(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$), entonces $\bar{\mathbf{x}}$ es una solución de $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t)$ que satisface $\mathbf{x}(t_0) = \mathbf{x}^0$. A las transformaciones de funciones en funciones se les llama a veces "operadores".

2.12 Definición. *Sea $\mathscr{F}$ un subconjunto de $\mathscr{C}$. Una transformación T de $\mathscr{F}$ en $\mathscr{C}$ se dice que es una* ***contracción*** *sobre $\mathscr{F}$ si*

$$\|T(\mathbf{x}) - T(\mathbf{y})\| \leqslant \lambda \|\mathbf{x} - \mathbf{y}\|$$

para todo $\mathbf{x}$, $\mathbf{y}$ *en $\mathscr{F}$ y algún $\lambda \in \langle 0, 1 \rangle$.*

2.13 Teorema. *Sea* $\mathbf{u}$ *una función cualquiera en $\mathscr{C}$. Definamos $\mathscr{F}$ como el conjunto de todas las funciones* $\mathbf{x}$ *en $\mathscr{C}$ que satisfacen $\|\mathbf{x} - \mathbf{u}\| \leqslant b$. Si T es una contracción sobre $\mathscr{F}$ y si T transforma $\mathscr{F}$ en sí mismo, entonces T tiene un punto único fijo $\bar{\mathbf{x}}$ en $\mathscr{F}$. En realidad, si $\mathbf{x}^0$ es una función cualquiera en $\mathscr{F}$, entonces $\mathbf{x}^n = T^{(n)}(\mathbf{x}^0) \to \bar{\mathbf{x}}$ cuando $n \to \infty$ y*

2.14 $$\|\bar{\mathbf{x}} - \mathbf{x}^n\| \leqslant \frac{\lambda}{1 - \lambda} \|\bar{\mathbf{x}}^n - \mathbf{x}^{n-1}\|.$$

PRUEBA. La prueba es esencialmente la misma que la del teorema 2.2. Reemplazamos $|x|$ por $\|\mathbf{x}\|$ y $\mathscr{I}$ por $\mathscr{F}$. La convergencia sobre la recta real se reemplaza por la convergencia uniforme de las funciones. En lugar del problema 2, pág. 480, tenemos el problema 13 que sigue. Necesitamos también observar que $\lim_{n \to \infty} \mathbf{y}^n = \mathbf{y}$ implica $\lim_{n \to \infty} \|\mathbf{y}^n\| = \|\mathbf{y}\|$. Esto es la continuidad de la norma. Por la desigualdad del triángulo

$$\|\mathbf{y}\| - \|\mathbf{y}^n - \mathbf{y}\| \leqslant \|\mathbf{y}^n\| \leqslant \|\mathbf{y} - \mathbf{y}^n\| + \|\mathbf{y}\|,$$

de donde se sigue que $\lim \|\mathbf{y}^n\| = \|\mathbf{y}\|$. La continuidad de la norma y el hecho de que el límite de una sucesión uniformemente convergente de funciones continuas es continua implica que $\mathscr{F}$ es cerrado ($\mathbf{y}^n$ en $\mathscr{F}$ y $\lim \mathbf{y}^n = \mathbf{y}$ implica que $\mathbf{y}$ está en $\mathscr{F}$). Además, vemos que como T es una contracción, luego continua: $\lim \|T(\mathbf{y}^n) - T(\mathbf{y})\| = 0$; es decir, $\lim T(\mathbf{y}^n) = T(\mathbf{y})$. La prueba del teorema 2.2 es, entonces, una copia exacta de la prueba de este teorema.

Problemas

1. Pruébese que *si T es una contracción sobre un intervalo $\mathscr{I}$, entonces T es continua en $\mathscr{I}$.*

2. Pruébese que *si T es diferenciable sobre un intervalo $\mathscr{I}$ y si* $|T'(x)| \leqslant \lambda < 1$ *para todo x en $\mathscr{I}$, entonces T es una contracción en $\mathscr{I}$.*

3. Pruébese que para $T(x) = \frac{x}{2} + \frac{1}{x}$ y para cualquier x_0 en $[1, 2]$, $T^{(n)}(x_0) \to \sqrt{2}$ cuando $n \to \infty$.

4. Bajo las condiciones del teorema 2.2 pruébese que

$$|\bar{x} - x_n| \leqslant \frac{\lambda^n}{1-\lambda} |T(x_0) - x_0|.$$

Nota. Esto demuestra la ventaja de comenzar con una buena aproximación inicial. Sin embargo, esto también prueba que tal cosa no es ni con mucho tan importante como arreglar el problema de forma que λ sea pequeño. Para resolver, por ejemplo, $f(x) = 0$ hay una infinidad de formas de convertir el problema en un problema equivalente de punto fijo. Algunos problemas de punto fijo darán una convergencia más rápida que otros de las aproximaciones sucesivas.

5. Sea $\mathscr{I}$ el intervalo cerrado definido por $|x - x_0| \leqslant a$ ($\mathscr{I} = [x_0 - a, x_0 + a]$). Pruébese que *si* 1) *T es una contracción sobre* $\mathscr{I} = [x_0 - a, x_0 + a]$ *con* $|T(x) - T(y)| \leqslant \lambda|x - y|$, $0 < \lambda < 1$ *para todo* x, y *en* $\mathscr{I}$, *y* 2) $|T(x_0) - x_0| \leqslant (1 - \lambda)a$, *entonces T tiene un punto fijo único* $\bar{x}$ *en* $\mathscr{I}$ *y* $T^{(n)}(x_0) \to \bar{x}$ *cuando* $n \to \infty$.

6. *El método de Newton.* Supongamos que

$$\left|\frac{f(x)\,f''(x)}{[f'(x)]^2}\right| \leqslant \lambda < 1 \quad \text{y} \quad \left|\frac{f(x_0)}{f'(x_0)}\right| \leqslant (1-\lambda)a$$

para todo x en $[x_0 - a, x_0 + a]$. Pruébese que la sucesión de aproximaciones sucesivas definida por la relación de recurrencia

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

converge a una solución de $f(x) = 0$.

7. Calcúlese con tres cifras significativas exactas la raíz positiva de $x - 2 \operatorname{sen} x = 0$.

8. Calcúlese con tres cifras significativas exactas la solución de $e^{-x} = x$.

9. Supongamos que T es continua sobre $[a, b]$ y que $a < T(a)$ y $T(b) < b$.

a) Conclúyase que hay puntos fijos de T en $\langle a, b \rangle$.

b) Si T es no creciente sobre $[a, b]$ pruébese que hay un punto fijo único de T en $\langle a, b \rangle$.

10. Pruébese que *a*) 2.5 *b*) 2.6 *c*) 2.7.

11. Generalícese el teorema 2.2 a transformaciones T de R^n en R^n.

12. *Teorema de función implícita.* Supongamos que 1) $f(x_0, y_0) = 0$, 2) f y $D_2 f$ son continuas en una vecindad de (x_0, y_0), 3) $D_2 f(x_0, y_0) \neq 0$. Usando el teorema 2.13, pruébese la existencia de una función φ continua sobre algún intervalo $[x_0 - \alpha, x_0 + \alpha]$, tal que $\varphi(x_0) = y_0$ y $f(x, \varphi(x)) = 0$ para todo x en $[x_0 - \alpha, x_0 + \alpha]$.

Sugerencia: consideremos $z = T(y)$ donde $z(x) = T(y)(x) = y(x) - \dfrac{f(x, y(x))}{D_2 f(x_0, y_0)}$.

13. Sea $\{\mathbf{x}^n\}$ una sucesión de funciones en $\mathscr{C}$ con la propiedad de que para algún $\lambda \in \langle 0, 1 \rangle$

$$\|\mathbf{x}^n - \mathbf{x}^{n-1}\| \leqslant \lambda \|\mathbf{x}^{n-1} - \mathbf{x}^{n-2}\|, \quad n \geqslant 2.$$

Pruébese que entonces la sucesión converge.

3. TEOREMA DE EXISTENCIA Y UNICIDAD PARA LAS ECUACIONES DIFERENCIALES

El primer teorema de existencia y unicidad para ecuaciones diferenciales lo probó alrededor de 1825 Agustín Cauchy (1789-1857), uno de los más grandes matemáticos de Francia. Unos cincuenta años más tarde, en 1876, R. Lipschitz (1832-1903) probó un teorema análogo bajo condiciones algo más débiles que las de Cauchy. El teorema que presentamos en esta sección usa la condición de Lipschitz, y el método de prueba, que es el de las aproximaciones sucesivas, lo introdujo por vez primera E. Picard (1857-1941) en 1890.

Consideramos un sistema

3.1
$$\begin{aligned} \dot{x}_1(t) &= f_1(x_1(t), x_2(t), \ldots, x_n(t), t) \\ \dot{x}_2(t) &= f_2(x_1(t), x_2(t), \ldots, x_n(t), t) \\ &\vdots \\ \dot{x}_n(t) &= f_n(x_1(t), x_2(t), \ldots, x_n(t), t), \end{aligned}$$

de n ecuaciones diferenciales de primer orden en n funciones incógnitas $x_1, x_2, \ldots, x_n$. La ecuación diferencial de orden n-ésimo

3.2
$$\frac{d^n x}{dt^n} = F\left(x, \frac{dx}{dt}, \cdots, \frac{d^{n-1} x}{dt^{n-1}}, t\right)$$

es un caso especial. Sea $x_1 = x$, $x_2 = \dfrac{dx}{dt}$, $\cdots$, $x_n = \dfrac{d^{n-1} x}{dt^n}$. Entonces 3.2

es equivalente al sistema

3.3
$$\begin{aligned}\dot{x}_1(t) &= x_2(t)\\ \dot{x}_2(t) &= x_3(t)\\ &\vdots\\ \dot{x}_{n-1}(t) &= x_n(t)\\ \dot{x}_n(t) &= F(x_1(t), x_2(t), \ldots, x_n(t), t).\end{aligned}$$

Si x es una solución de 3.2, entonces $(x_1, x_2, \cdots, x_n) = \left(x, \frac{dx}{dt}, \cdots, \frac{d^{n-1}x}{dt^{n-1}}\right)$ es una solución de 3.3. Recíprocamente, si $(x_1, \ldots, x_n)$ es una solución de 3.3, entonces x_1 es una solución de 3.2.

En notación vectorial 3.1 puede escribirse

3.4
$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t),$$

donde $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{f} = (f_1, \ldots, f_n)$. La función $\mathbf{f}$ es una función de $R^n \times R$ en R^n.

Sea $\mathscr{I}$ el intervalo cerrado $[t_0 - a, t_0 + a]$, $\mathscr{S} = \{\mathbf{v} \in R^n \mid |\mathbf{v} - \mathbf{v}^0| \leqslant b\}$, y $\mathscr{E} = \mathscr{S} \times \mathscr{I} = \{(\mathbf{v}, t) \mid \mathbf{v} \in \mathscr{S}, t \in \mathscr{I}\}$. El conjunto $\mathscr{E}$ es un conjunto cerrado y acotado de puntos en R^{n+1}. Como antes $\mathscr{C}$ denotará el conjunto de funciones continuas $\mathbf{x}$ de $\mathscr{S}$ en R^n. $\mathscr{F}$ será el conjunto de todas las funciones continuas $\mathbf{x}$ de $\mathscr{I}$ a $\mathscr{S}$. El conjunto $\mathscr{F}$ es el subconjunto de $\mathscr{C}$ con la propiedad de que $\mathbf{x}$ en $\mathscr{F}$ implica $(\mathbf{x}(t), t)$ en $\mathscr{E}$ para todo t en $\mathscr{I}$ (la gráfica de $\mathbf{x}$ está en $\mathscr{E}$).

Supongamos que $\mathbf{f}$ es continua sobre $\mathscr{E}$. Entonces la transformación T definida por

3.5
$$\mathbf{y}(t) = T(\mathbf{x})(t) = \mathbf{v}^0 + \int_{t_0}^{t} \mathbf{f}(\mathbf{x}(\tau), \tau)\, d\tau$$

está definida sobre $\mathscr{F}$. Vemos entonces que $\mathbf{x}$ *está en* $\mathscr{F}$ *es una solución de* 3.4 *que satisface* $\mathbf{x}(t_0) = \mathbf{v}^0$ *si y sólo si* $\mathbf{x}$ *es un punto fijo de* T. Si $\mathbf{x}$ es un punto fijo ($T(\mathbf{x}) = \mathbf{x}$) en $\mathscr{F}$, entonces

3.6
$$\mathbf{x}(t) = \mathbf{v}^0 + \int_{t_0}^{t} \mathbf{f}(\mathbf{x}(\tau), \tau)\, d\tau, \quad t \in \mathscr{I},$$

y

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t), \quad t \in \mathscr{I}$$

con $\mathbf{x}(t_0) = \mathbf{v}^0$. Recíprocamente, si $\mathbf{x}$ es una solución de 3.4 que satisface $\mathbf{x}(t_0) = \mathbf{v}^0$, entonces obtenemos 3.6 por integración. Estamos ahora en posición de aplicar nuestro teorema de punto fijo para transformaciones de funciones.

3.7 Definición. *La función* $\mathbf{f}$ *se dice que satisface una* ***condición de Lipschitz*** *en* $\mathscr{E}$ *con respecto a* $\mathbf{v}$ *si hay un número real* K *con la propiedad de que*

$$|\mathbf{f}(\mathbf{v}^1, t)-\mathbf{f}(\mathbf{v}^2, t)| \leqslant K|\mathbf{v}^1-\mathbf{v}^2|$$

para cualesquier $(\mathbf{v}^1, t)$, $(\mathbf{v}^2, t)$ *en* $\mathscr{E}$.

La condición de Cauchy fue la hipótesis de que $\mathbf{f}$ era continua sobre $\mathscr{E}$ y tenía primeras derivadas parciales continuas $D_1\mathbf{f}, \ldots, D_n\mathbf{f}$ sobre $\mathscr{E}$. Para la mayor parte de las aplicaciones puede admitirse la hipótesis de Cauchy, y esta condición implica la condición de Lipschitz (problema 5).

3.8 Teorema. *Si la función* $\mathbf{f}$ *es continua y satisface una condición de Lipschitz (con respecto a* $\mathbf{v}$*) en una vecindad* $\mathscr{E}_1$ *de* $(\mathbf{v}^0, t)$, *entonces para algún* $a>0$ *hay una solución única sobre* $\mathscr{J}=[t_0-a, t_0+a]$ *de* $\dot{\mathbf{x}}(t)=\mathbf{f}(\mathbf{x}(t), t)$ *que satisface* $\mathbf{x}(t_0)=\mathbf{v}^0$.

Prueba. Nuestra hipótesis es (expresada con la notación que antes hemos adoptado) que $\mathbf{f}$ satisface una condición de Lipschitz

$$|\mathbf{f}(\mathbf{v}^1, t)-\mathbf{f}(\mathbf{v}^2, t)| \leqslant K|\mathbf{v}^1-\mathbf{v}^2|$$

para todo $(\mathbf{v}^1, t)$ y $(\mathbf{v}^2, t)$ en $\mathscr{E}_1=\{(\mathbf{v}, t) \mid |\mathbf{v}-\mathbf{v}^0|<b, |t-t_0|<a_1\}=\mathscr{S}\times\mathscr{J}_1$, y que $\mathbf{f}$ es continua en $\mathscr{E}_1$. La prueba consistirá en demostrar que para una a suficientemente pequeña la transformación T definida por 3.5 es una contracción sobre $\mathscr{F}$ que transforma $\mathscr{F}$ en sí mismo. Entonces, por el teorema 2.13, T tiene un punto fijo único en $\mathscr{F}$, que es entonces la solución única que estamos buscando. Demostramos primero que para a suficientemente pequeña $T(\mathscr{F})\subset\mathscr{F}$. Sea Máx $\{|\mathbf{f}(\mathbf{v}, t)| \mid (\mathbf{v}, t)\in\mathscr{E}_1\}=M$. Para todo $\mathbf{x}$ en $\mathscr{F}_1$, $\mathbf{y}=T(\mathbf{x})$ satisface

$$|\mathbf{y}(t)-\mathbf{v}^0|=\left|\int_{t_0}^{t}\mathbf{f}(\mathbf{x}(\tau), \tau)\,d\tau\right| \leqslant \left|\int_{t_0}^{t}|\mathbf{f}(\mathbf{x}(\tau), \tau)|\,d\tau\right|$$

$$\leqslant M|t-t_0| \leqslant Ma_1$$

para todo $t\in\mathscr{J}_1$. Elíjase a de modo que $a\leqslant a_1$ y $Ma\leqslant b$. Entonces, recordando la definición de la norma como el valor máximo, tenemos

$$\|\mathbf{y}-\mathbf{v}^0\| = \|T(\mathbf{x})-\mathbf{v}^0\| \leqslant b$$

para todo $\mathbf{x}$ en $\mathscr{F}$. Esto significa para un a suficientemente pequeño que $T(\mathscr{F})\subset\mathscr{F}$, y T transforma $\mathscr{F}$ en sí mismo. Queda por demostrar que para a suficientemente pequeño T es una contracción sobre $\mathscr{F}$. Aquí es

donde necesitamos la condición de Lipschitz. Sean $\mathbf{y}^1 = T(\mathbf{x}^1)$ y $\mathbf{y}^2 = T(\mathbf{x}^2)$ donde $\mathbf{x}^1$ y $\mathbf{x}^2$ están en $\mathscr{F}$. Entonces

$$\begin{aligned}
|\mathbf{y}^1(t)-\mathbf{y}^2(t)| &= \left|\int_{t_0}^{t} [\mathbf{f}(\mathbf{x}^1(\tau),\tau)-\mathbf{f}(\mathbf{x}^2(\tau),\tau)]\,d\tau\right| \\
&\leqslant \left|\int_{t_0}^{t} |\mathbf{f}(\mathbf{x}^1(\tau),\tau)-\mathbf{f}(\mathbf{x}^2(\tau),\tau)|\,d\tau\right| \\
&\leqslant K\left|\int_{t_0}^{t} |\mathbf{x}^1(\tau)-\mathbf{x}^2(\tau)|\,d\tau\right| \\
&\leqslant K\left|\int_{t_0}^{t} \|\mathbf{x}^1-\mathbf{x}^2\|\,d\tau\right| \\
&= K\,|t-t_0|\,\|\mathbf{x}^1-\mathbf{x}^2\|.
\end{aligned}$$

Por tanto, para $\mathbf{x}^1$, $\mathbf{x}^2$ en $\mathscr{F}$

$$\|T(\mathbf{x}^1)-T(\mathbf{x}^2)\| \leqslant aK\|\mathbf{x}^1-\mathbf{x}^2\|.$$

Elijamos a de forma que $aK < 1$. Entonces T es una contracción sobre $\mathscr{F}$. Y esto completa la prueba.

Como una consecuencia del teorema 2.13 sabemos también que para a y b suficientemente pequeños y cualquier $\mathbf{x}^0$ en $\mathscr{F}$, la sucesión $\mathbf{x}^n = T(\mathbf{x}^{n-1}) = T^{(n)}(\mathbf{x}^0)$ converge uniformemente sobre $\mathscr{I}$ a la solución de $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t)$ que satisface $\mathbf{x}(t_0) = \mathbf{v}^0$. Por ejemplo, siempre se puede tomar $\mathbf{x}^0$ como la función definida por $\mathbf{x}^0(t) = \mathbf{v}^0$. Aunque éste no es a menudo un procedimiento práctico para calcular una solución de una ecuación diferencial, es un método de gran generalidad y tiene otras aplicaciones. Consideraremos dos ejemplos, uno extremadamente sencillo para ilustrar el método y el otro para ilustrar la dificultad.

3.9 Ejemplo. Obtener aproximaciones sucesivas de

a) $\dot{x}(t) = 2x(t)-4t$, $x(0) = 0$

b) $\dot{x} = \sqrt{1+x^2}$, $x(0) = 1$.

SOLUCIÓN. *a*) $T(x)(t) = \int_0^t (2x(\tau)-4\tau)\,d\tau = -2t^2+2\int_0^t x(\tau)\,d\tau$. Tomando $x^0 = 0$, tenemos

$$\begin{aligned}
x^1(t) &= T(x^0)(t) = -2t^2, \\
x^2(t) &= T(x^1)(t) = -2t^2-4\int_0^t \tau^2\,d\tau \\
&= -2t^2-\tfrac{4}{3}t^3,
\end{aligned}$$

$$x^3(t) = T(x^2)(t) = -2t^2-4\int_0^t \left(\tau^2 + \frac{2}{3}\tau^3\right)d\tau$$

$$= -2t^2 - \frac{4}{3}t^3 - \frac{8}{3\cdot 4}t^4$$

$$x^n(t) = -\left(\frac{(2t)^2}{2!} + \frac{(2t)^3}{3!} + \cdots + \frac{(2t)^{n+1}}{(n+1)!}\right).$$

Aquí vemos que la sucesión de aproximaciones sucesivas converge a $x(t) = 1+2t-e^{2t}$, que es la solución que podríamos haber obtenido directamente resolviendo la ecuación lineal.

b) $T(x)(t) = 1 + \int_0^t \sqrt{1+x^2(\tau)}\,d\tau$. Hagamos $x^0 = 1$. Entonces

$$x^1(t) = T(x^0)(t) = 1 + \int_0^t \sqrt{2}\,d\tau = 1+\sqrt{2}t$$

$$x^2(t) = T(x^1)(t) = 1 + \int_0^t \sqrt{1+(1+\sqrt{2}\tau)^2}\,d\tau$$

$$= \frac{1}{2} - \frac{1}{2\sqrt{2}}\ln(1+\sqrt{2}) + \frac{1}{2\sqrt{2}}\left[(1+\sqrt{2}t)\sqrt{1+(1+\sqrt{2}t)^2}\right.$$

$$\left. + \ln(1+\sqrt{2}t + \sqrt{1+(1+\sqrt{2}t)^2})\right].$$

Ya a esta altura podemos darnos cuenta de la dificultad de seguir calculando aproximaciones sucesivas.

El valor de este método de aproximaciones sucesivas en las ecuaciones diferenciales no estriba solamente en que él nos da un teorema de existencia y unicidad, sino que puede también usarse para derivar propiedades generales de las soluciones: la dependencia de las soluciones de las condiciones iniciales, la dependencia de las soluciones de un parámetro que aparece en las ecuaciones diferenciales, y así por el estilo. Usualmente no es un método práctico para el cálculo numérico de las soluciones.

Pensemos de nuevo en el teorema de existencia y unicidad. Nótese primero que es un teorema local. Solamente afirma la existencia y unicidad de una solución en una vecindad del punto inicial $(\mathbf{v}^0, t_0)$. La solución se sabe que existe para todo t que satisfaga $|t-t_0| \leqslant a$, donde a es suficientemente pequeño. La prueba nos permite hacer una estimación de a que depende de la cuantía de K y M. Sin embargo, puede que sea posible continuar la

solución mucho más de lo que esta estimación indica. Supongamos que las condiciones del teorema 3.8 se satisfacen en una vecindad de cada punto $(\mathbf{v}, t)$ de un conjunto abierto $\mathscr{D}$ de $R^n \times R$. Consideremos la solución $\mathbf{x}$ definida en una vecindad de un punto $(\mathbf{v}^0, t)$ en $\mathscr{D}$ que pasa por este punto (es decir, tal que $\mathbf{x}(t_0) = \mathbf{v}^0$). Sabemos que la solución está definida sobre un intervalo $[t_0 - a, t_0 + a]$ y permanece en $\mathscr{D}$. Hay, por tanto, una solución única a través del punto $(\mathbf{x}(t_0 + a), t_0 + a)$. A la izquierda de $t_0 + a$ coincide con la solución con la que hemos comenzado. Esta es, por tanto, una extensión única de la solución $\overline{\mathbf{x}}$. Continuando en esta forma tanto para t creciente como para t decreciente, puede argumentarse que hay un intervalo máximo $\mathscr{I} = \langle \alpha, \beta \rangle$ sobre el que la solución puede continuarse en $\mathscr{D}$. Esta solución sobre $\mathscr{I}$ se llama *solución completa* en $\mathscr{D}$ que pasa por el punto $(\mathbf{v}^0, t_0)$. Es, en general, difícil determinar la amplitud del intervalo $\mathscr{I}$. Puede, por ejemplo, ser finito. La ecuación no líneal más sencilla ilustra lo que puede suceder. La solución completa de $\dot{x} = x^2$ que satisface $x(0) = 1$ es $(1-t)^{-1}$. El intervalo máximo $\mathscr{I}$ es $\langle -\infty, 1 \rangle$. La solución va a infinito en un tiempo finito y no puede extenderse más allá de $t = 1$. Para el sistema lineal $\dot{\mathbf{x}}(t) = A(t)\,\mathbf{x}(t) + \mathbf{f}(t)$ probamos para coeficientes constantes y supusimos cierto para no constantes que, si A y $\mathbf{f}$ son continuos sobre $\langle -\infty, \infty \rangle$, entonces las soluciones completas están definidas sobre $\langle -\infty, \infty \rangle$. Para sistemas no lineales no hay respuestas tan sencillas.

Nota. Todo lo que hemos probado en esta sección y la sección previa puede extenderse fácilmente a funciones vectoriales complejas y la existencia y unicidad se aplica a la ecuación diferencial vectorial compleja $\dot{\mathbf{z}} = \mathbf{f}(\mathbf{z}, t)$. La función $\mathbf{f}$ es de $C^n \times R$ en C^n, y una solución es una función $\mathbf{z}$ de R en C^n. Es suficiente reemplazar C^n por R^{2n}.

Problemas

1. Úsese el método de aproximaciones sucesivas para resolver

a) $\dot{x}(t) + 2x(t) = 3t^2 - 1,\ x(0) = 1$
b) $\ddot{x} + x = 0,\ x(0) = 0,\ \dot{x}(0) = 1$
c) $2\dot{x}(t) + tx(t) = 0,\ x(0) = 1.$

2. Obténgase, comenzando con la condición inicial, las tres primeras aproximaciones sucesivas de

a) $\dot{x} = x^2,\ x(0) = 1$

b) $\dfrac{dy}{dx} = x^2 + y^2,\quad y(0) = 1$

c) $\dot{x} = x^2 + y^2$
$\dot{y} = x - y,\ x(0) = 0,\ y(0) = 1$
d) $\ddot{x} + x - \frac{1}{3}x^3 = 0,\ x(0) = \dot{x}(0) = 1.$

3. Considérese la ecuación diferencial autónoma (**f** es independiente de t)

$$(*) \qquad \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t))$$

donde **f** es una función de R^n en R^n que es continua sobre un conjunto abierto $\mathscr{E}$ en R^n. Supongamos que **x** es una solución de (*) sobre $[0, \infty\rangle$ y que $\mathbf{x}(t) \to \mathbf{v}^0$ cuando $t \to \infty$ donde $\mathbf{v}^0$ es un punto de $\mathscr{E}$. Conclúyase que $\mathbf{v}^0$ es un estado de equilibrio de (*); es decir, $\mathbf{f}(\mathbf{v}^0) = \mathbf{0}$.

4. Sea $F(x, y)$ una función de R^2 en R con derivadas parciales continuas de segundo orden sobre un conjunto abierto $\mathscr{E}$ de R^2. Sea (x_0, y_0) un punto en $\mathscr{E}$ y supongamos que $[D_1 F(x_0, y_0)]^2 + [D_2 F(x_0, y_0)]^2 \neq 0$. Pruébese que hay entonces una curva lisa γ que pasa por (x_0, y_0) sobre la curva de nivel $F(x, y) = F(x_0, y_0)$. Si $D_2 F(x_0, y_0) \neq 0$, pruébese que γ puede representarse en la vecindad de (x_0, y_0) como la gráfica de una función $y = f(x)$. Formúlense estos resultados como teoremas de función implícita.

5. Sea **f** una función de R^{n+1} a R^n continua sobre el conjunto $\mathscr{E} = \mathscr{S} \times \mathscr{I}$ donde $\mathscr{S} = \{\mathbf{v} \mid |\mathbf{v} - \mathbf{v}^0| \leqslant b\} \subset R^n$ e $\mathscr{I} = [t_0 - a, t_0 + a]$ y sean $D_1 \mathbf{f}, \ldots, D_n \mathbf{f}$ continuas sobre $\mathscr{E}$. Demuéstrese que **f** satisface la condición de Lipschitz: hay un número real K tal que

$$|\mathbf{f}(\mathbf{v}^1, t) - \mathbf{f}(\mathbf{v}^2, t)| \leqslant K|\mathbf{v}^1 - \mathbf{v}^2|$$

para cualesquier $(\mathbf{v}^1, t)$, $(\mathbf{v}^2, t)$ en $\mathscr{E}$.

6. Considérese la ecuación diferencial autónoma

$$(*) \qquad \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t))$$

donde **f** es continua y satisface una condición de Lipschitz en la vecindad de cada punto $\mathbf{v}^0$ de R^n. Sea **x** una solución de (*). Pruébese que **x** es periódica si y sólo si $\mathbf{x}(t_1) = \mathbf{x}(t_2)$ para algún $t_1 \neq t_2$ (es decir, la curva descrita por la solución es cerrada). ¿Qué es el periodo mínimo?

4. FUNCIONES CIRCULARES

Ahora que tenemos un teorema de existencia y unicidad para las ecuaciones diferenciales, sabemos que ecuaciones diferenciales más condiciones iniciales definen funciones. Éste es el origen de muchas de las funciones cuyas propiedades han sido estudiadas por los matemáticos y para las cuales se han calculado tablas. Como un sencillo ejemplo de esto definiremos las funciones circulares por medio de una ecuación diferencial.

Puede ser que las funciones circulares o trigonométricas se hayan definido bien en el curso de estudios del lector, o puede que no. Pero cualquiera que sea el caso nosotros vamos a imaginar en lo que sigue que

nada sabemos de estas funciones, y que estamos interesados en las soluciones de la ecuación diferencial

4.1 $$\ddot{x}+x = 0.$$

Un sistema equivalente es

4.2 $$\begin{aligned} \dot{x} &= -y \\ \dot{y} &= x, \end{aligned}$$

y éste puede reemplazarse por una sola ecuación compleja por la sustitución $z = x+iy$. Entonces $\dot{z} = \dot{x}+i\dot{y} = -y+ix = iz$, de modo que tenemos que habérnosla con la ecuación diferencial

$$\dot{z} = iz.$$

Es ventajoso y no más difícil, estudiar la ecuación diferencial

4.3 $$\dot{z} = \lambda z,$$

donde λ es un número complejo arbitrario. Designemos por E la solución de 4.3 que satisface $z(0) = 1$. Por el teorema de existencia y unicidad extendido a funciones complejas, sabemos que E está definida al menos en una vecindad de 0. Una de las cosas que necesitamos probar primero es que E está definida sobre $\langle -\infty, \infty\rangle$. La función E es un punto fijo (el único punto fijo) de

$$T(z)(t) = 1+\lambda \int_0^t z(\tau)\,d\tau,$$

y es el límite de una sucesión de aproximaciones sucesivas.

Tomando $z_0 = 1$, la condición inicial, obtenemos

$$z_1(t) = T(z_0)(t) = 1+\lambda \int_0^t z_0(\tau)\,d\tau = 1+\lambda t,$$

$$z_2(t) = T(z_1)(t) = 1+\lambda \int_0^t (1+\lambda\tau)\,d\tau = 1+\lambda t + \frac{(\lambda t)^2}{2!}$$

$$z_3(t) = 1+\lambda t + \frac{(\lambda t)^2}{2!} + \frac{(\lambda t)^3}{3!}$$

y, por inducción,

$$z_n(t) = 1+\lambda t + \cdots + \frac{(\lambda t)^n}{n!} = \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!}$$

La sucesión $\{z_n\}$ converge absoluta y uniformemente sobre todo intervalo

finito. De aquí se sigue (o puede verificarse directamente por sustitución en 4.3) que

4.4 $$E(t) = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!}$$

es la solución sobre $\langle -\infty, \infty \rangle$ de 4.3 que satisface $z(0) = 1$. Si λ es un número real, entonces $E(t) = e^{\lambda t}$, y si λ es un número complejo $E(t)$ es entonces la extensión de la función exponencial real. Escribimos para toda λ real o compleja, $E(t) = e^{\lambda t}$. De donde

$$\frac{d}{dt} e^{\lambda t} = \lambda e^{\lambda t} \quad \text{y} \quad e^0 = 1 .$$

Esta es la única función con estas propiedades.

Usemos ahora la ecuación diferencial 4.3 para derivar algunas propiedades de $e^{\lambda t}$. Sabemos, por el teorema de unicidad, que $ce^{\lambda t}$ es la única solución sobre $\langle -\infty, \infty \rangle$ de 4.3 que satisface $z(0) = c$. Consideremos ahora las funciones

$$E_1(t) = e^{\alpha} e^{\lambda t} \quad \text{y} \quad E_2(t) = e^{\alpha + \lambda t}.$$

Como $E_1'(t) = \lambda E_1(t)$ y $E_2'(t) = \lambda E_2(t)$, E_1 y E_2 son, ambas, soluciones de 4.3. Satisfacen las mismas condiciones iniciales $E_1(0) = E_2(0) = e^{\alpha}$ y, por tanto, por la unicidad de las soluciones, $E_1(t) = E_2(t)$ para todo t. Tomando $t = 1$ tenemos la ley de los exponentes

4.5 $$e^{\alpha} e^{\lambda} = e^{\alpha + \lambda}.$$

En particular, $e^{\lambda} e^{-\lambda} = e^0 = 1$ y, por tanto,

$$e^{-\lambda} = (e^{\lambda})^{-1}.$$

Vemos inmediatamente por esto que $e^{\lambda} \neq 0$ para cualquier número complejo λ.

En el capítulo 11, sección 4, usamos las funciones seno y coseno para definir e^{it}. Ahora que ya hemos definido e^{it} independientemente de las funciones circulares, podemos seguir el camino inverso y usar esta función exponencial compleja para definir el seno y el coseno. Definimos

4.6 $$\cos t = \tfrac{1}{2}(e^{it} + e^{-it}), \quad \operatorname{sen} t = \frac{1}{2i}(e^{it} - e^{-it}).$$

Así pues, $e^{it} = \cos t + i \operatorname{sen} t$ y la función vectorial (cos, sen) es entonces la única solución de 4.2 que satisface $(x(0), y(0)) = (1, 0)$.

Podemos ahora derivar las propiedades de las funciones seno y coseno

de las propiedades de e^{it} o directamente de la ecuación diferencial 4.2. Por 4.4 sabemos que la serie para e^{it} es

$$e^{it} = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} = 1 + it - \frac{t^2}{2!} - i\frac{t^3}{3!} + \frac{t^4}{4!} + \cdots$$

y por tanto

$$\cos t = 1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!}$$

$$\operatorname{sen} t = t - \frac{t^3}{3!} + \frac{t^5}{5!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!}.$$

Como $(x, y) = (\cos, \operatorname{sen})$ es una solución de 4.2 vemos que

$$D \cos = -\operatorname{sen} \quad \text{y} \quad D \operatorname{sen} = \cos.$$

La curva γ definida por $x = \cos t$, $y = \operatorname{sen} t$ se encuentra sobre la curva integral $x^2 + y^2 = 1$ de 4.2. Así pues, $\cos^2 + \operatorname{sen}^2 = 1$.

Probamos ahora que seno y coseno tienen 2π como periodo mínimo. No es difícil demostrar que hay un tiempo mínimo $T > 0$ con la propiedad de que $\cos T = 0$ y $\operatorname{sen} T = 1$. Como (cos, sen) es una solución de 4.2 con la condición inicial $(\cos 0, \operatorname{sen} 0) = (1, 0)$, $\cos t > 0$ para todo t positivo y suficientemente pequeño. Para tales t la curva γ está en el primer cuadrante, ya que de 4.2 se sigue que $\operatorname{sen} t$ es creciente y $\cos t$ es decreciente. Además, después de un tiempo $t > \varepsilon > 0$, $\dot{x} = -y < -\operatorname{sen} \varepsilon < 0$. Así pues, en un tiempo finito la curva γ alcanza el eje Y y x se hace negativo. Por tanto, hay un $T > 0$ mínimo con la propiedad de que $\cos T = 0$ y $\operatorname{sen} T = 1$. Como $x^2 + y^2 = 1$ sobre γ, t es la longitud del arco a lo largo de la circunferencia (unitaria). Se sigue entonces fácilmente que en el tiempo $4T$ el punto $x = \cos t$, $y = \operatorname{sen} t$ ha dado una vuelta alrededor de la circunferencia y $4T = 2\pi$. Vemos, por tanto, que 2π es el periodo mínimo del seno y el coseno. Como $e^{it} = \cos t + i \operatorname{sen} t$, 2π es también el periodo mínimo de e^{it}.

Hemos, pues, definido las funciones circulares seno y coseno, hemos hallado su expansión en series de potencias para calcularlas y hemos deducido todas las propiedades básicas de estas funciones. Las propiedades de adición del seno y el coseno se siguen fácilmente de la ley de exponentes 4.5 como se mostró en la sección 4 del capítulo 11.

5. SOLUCIÓN EN SERIE DE LAS ECUACIONES DIFERENCIALES

La idea de obtener soluciones en serie de las ecuaciones diferenciales se remonta más de 300 años hasta Newton. Una de las ecuaciones estudiadas

por Newton era

5.1
$$\frac{dy}{dx} = 1 + \frac{y}{1-x}, \quad y(0) = 0.$$

Desarrollando $(1-x)^{-1}$ escribió la ecuación en la forma

$$\frac{dy}{dx} = 1 + y(1 + x + x^2 + \cdots)$$

y obtuvo la solución

$$y(x) = x + \tfrac{1}{2}(x^2 + x^3 + \ldots).$$

Veremos dentro de un momento cómo es que esto puede hacerse. Notando que $x^2 + x^3 + \ldots = x^2/(1-x)$, la solución es

$$y(x) = x + \tfrac{1}{2}x^2/(1-x).$$

En lo que estamos interesados es en los métodos para la obtención de aproximaciones en serie y, con unas pocas palabras de advertencia, ésta que hemos expuesto es adecuada para ser la primera que consideremos. La advertencia es que este ejemplo no debe llevarnos a creer que va a sernos siempre posible, y especialmente con las ecuaciones no lineales, obtener la serie completa, e incluso cuando esto es posible no se deben tener demasiadas esperanzas en la posibilidad de expresar la suma de la serie en términos de funciones conocidas. Para muchos propósitos todo lo que se necesita es obtener sólo unos cuantos de los primeros términos y para muchas aplicaciones esto nos da una útil primera aproximación. Tal aproximación es a menudo, como más tarde veremos, el punto de partida para un cálculo numérico de la solución.

Ilustramos ahora dos métodos para obtener aproximaciones en serie. Supongamos que 5.1 tiene una solución analítica en la vecindad del origen:

5.2
$$y(x) = a_0 + a_1 x + a_1 x^2 + \cdots = \sum_{n=0}^{\infty} a_n x^n.$$

Método 1. Serie de Taylor

Sabemos que los coeficientes a_n vienen dados por $n!a_n = y^{(n)}(0)$. Por la ecuación diferencial 5.1 podemos entonces calcular sucesivamente tantas derivadas evaluadas en el origen como deseemos.

$$\begin{aligned}
y(0) &= 0 = a_0\\
y'(0) &= 1 + y(0) = 1 = a_1\\
y''(x) &= (1-x)^{-1}y'(x) + (1-x)^{-2}y(x)\\
y''(0) &= 1 = 2a_2\\
y'''(x) &= (1-x)^{-1}y''(x) + 2(1-x)^{-2}y'(x) + 2(1-x)^{-3}y(x)\\
y'''(0) &= 3 = 3!a_3.
\end{aligned}$$

Tenemos entonces como una solución aproximada

$$y(x) = x+\tfrac{1}{2}x^2+\tfrac{1}{2}x^3.$$

Para esta ecuación en particular no sería muy difícil probar que $y^{(n)}(0) = \frac{1}{2}n!$ y que $a_n = \frac{1}{2}$.

Método 2. Coeficientes indeterminados

Antes de comenzar a exponer este método, recuérdense nuestros resultados sobre diferenciación, adición e igualdad de series de potencias (capítulo 9, sección 9).

Es conveniente escribir la ecuación diferencial 5.1 en la forma

$$(1-x)\frac{dy}{dx} = 1-x+y.$$

Supongamos, ante todo, que estamos interesados en una aproximación hasta el orden 3. Los puntos denotan los términos de orden más alto. Sustituyendo la serie 5.2 en esta ecuación diferencial, obtenemos

$$\begin{aligned}(1-x)(a_1+2a_2x+3a_3x^2+4a_4x^3+\ldots)\\ = 1-x+a_0+a_1x+a_2x^2+a_3x^3+\ldots\end{aligned}$$

Igualando los coeficientes de potencias iguales de x, tenemos las ecuaciones

$$\begin{aligned}a_1 &= 1+a_0\\ 2a_2-a_1 &= -1+a_1\\ 3a_3-2a_2 &= a_2.\end{aligned}$$

Como $a_0 = y(0) = 0$, obtenemos

$$\begin{aligned}a_1 &= 1\\ a_2 &= -\tfrac{1}{2}+a_1 = \tfrac{1}{2}\end{aligned}$$

y la aproximación deseada es

$$y(x) = x+\tfrac{1}{2}x^2+\tfrac{1}{2}x^3.$$

Supongamos ahora que deseamos ver si es posible encontrar el coeficiente general a_n. Tenemos entonces al sustituir $y = \sum_{n=0}^{\infty} a_n x^n$ en la ecuación

$$(1-x)\left(\sum_{n=0}^{\infty} na_nx^{n-1}\right) = 1-x+\sum_{n=0}^{\infty} a_nx^n.$$

Agrupando términos, obtenemos

$$\sum_{n=0}^{\infty} [(n+1)a_{n+1}-na_n-a_n]x^n = 1-x.$$

De donde

$$a_1 - a_0 = 1$$
$$2a_2 - 2a_1 = -1$$
$$(n+1)(a_{n+1} - a_n) = 0, \quad n \geqslant 2.$$

Este sistema infinito de ecuaciones lineales se resuelve fácilmente

$$a_1 = 1 + a_0 = 1$$
$$a_2 = -\tfrac{1}{2} + a_1 = \tfrac{1}{2}$$
$$a_{n+1} = a_n = \tfrac{1}{2}, \quad n \geqslant 2.$$

Por tanto

$$y(x) = x + \tfrac{1}{2} \sum_{n=2}^{\infty} x^n, \quad |x| < 1.$$

El intervalo $\langle -1, 1\rangle$ es el intervalo de convergencia de la serie. Realmente, sabemos que

$$y(x) = x + \frac{1}{2} \frac{x^2}{1-x}, \quad x \neq 1.$$

El mayor intervalo en el que la solución que satisface $y(0) = 0$ está definida es $\langle -\infty, 1\rangle$. La solución aproximada, hasta el término de orden 3, es, según sabemos, suficientemente buena para x pequeño y no muy buena cuando x es cercano a 1.

Para una discusión más detallada de las aproximaciones en serie, para una justificación de la hipótesis de que hay una solución analítica, para el estudio de estimaciones del error, enviamos al lector a las referencias enumeradas en la bibliografía, particularmente a [58].

En el próximo ejemplo ilustramos la aplicación de estos métodos a una ecuación de segundo grado.

5.3 Ejemplo. Obténganse aproximaciones en serie de la solución de

5.4 $$u'' + \operatorname{sen} u = 0, \quad u(0) = 0, \quad u'(0) = u_0.$$

Ésta es la ecuación que describe el movimiento de un péndulo (ejemplo 14.8, pág. 688).

SOLUCIONES.

Método I. Supongamos una solución de la forma

$$u(t) = a_0 + a_1 t + a_2 t + \cdots$$

Entonces

$$u(0) = 0 = a_0, \; u'(0) = u_0 = a_1$$
$$u''(0) = -\operatorname{sen} u(0) = 0 = 2!\, a_2$$

$$u'''(0) = -u'(0)\cos u(0) = -u_0 = 3!\,a_3$$
$$u^{(4)}(0) = -u''(0)\cos u(0) + (u'(0))^2 \operatorname{sen} u(0) = 0 = 4!\,a_4$$
$$u^{(5)}(0) = -u'''(0)\cos u(0) + 3u''(0)\,u'(0)\operatorname{sen} u(0)$$
$$+ (u'(0))^3 \cos u(0) = u_0 + u_0^{\,3} = 5!\,a_5\,.$$

De donde

$$u(t) = u_0 t - \frac{1}{3!}u_0 t^3 + \frac{1}{5!}u_0(1+u_0^{\,2})t^5 + \cdots$$

Método II. Puede conseguirse una simplificación considerable si notamos de inmediato que la solución es impar (problema 6, pág. 537). Podemos entonces suponer una serie solución de la forma

$$u(t) = a_1 t + a_3 t^3 + a_5 t^5 + \cdots$$

Este uso de la simetría corta en dos el número de coeficientes a determinar. Reemplazando sen u por su serie y sustituyendo la serie para $u(t)$ en 5.4, tenemos

$$3\cdot 2a_3 t + 5\cdot 4a_5 t^3 + \cdots + (a_1 t + a_3 t^3 + \cdots)$$

$$-\frac{1}{3!}(a_1 t + a_3 t^3 + \cdots)^3 + \cdots = 0,$$

donde los puntos representan términos de orden mayor que 3. Igualando los coeficientes de potencias iguales de t obtenemos

$$6a_3 + a_1 = 0$$

$$20a_5 + a_3 - \frac{1}{3!}a_1^{\,3} = 0.$$

Como

$$u'(0) = a_1 = u_0,$$
$$a_3 = -\tfrac{1}{6}u_0$$
$$a_5 = \tfrac{1}{20}(\tfrac{1}{6}u_0 + \tfrac{1}{6}u^3).$$

Una solución aproximada es

$$u(t) = u_0 t - \frac{u_0}{3!}t^3 + \frac{u_0(1+u_0^{\,2})}{5!}t^5\,.$$

La ecuación para el siguiente coeficiente es

$$7\cdot 6a_7 + a_5 - \frac{3}{3!}a_1^{\,2}a_3 + \frac{1}{5!}a_1^{\,5} = 0,$$

y con esto podemos comenzar a ver la dificultad para determinar el término general.

Problemas

1. Determínense los primeros cuatro términos distintos de cero de la expansión en serie de la solución de

$$\dot{x}(t) = \text{sen}\, x(t) + e^t, \qquad x(0) = \pi$$

por *a*) método I; *b*) método II.

2. Úsese el método I para determinar la solución en serie de

$$\dot{x} = x^2, \quad x(0) = 1.$$

Sugerencia: para obtener el término general úsese la regla de Leibniz para la derivada n-ésima de un producto:

$$D^n(fg) = \sum_{k=0}^{n} \binom{n}{k} D^{n-k} f D^k g.$$

3. En el ejemplo 5.3 pruébese que $u_0 = \dot{\theta}_0 \sqrt{\frac{l}{g}}$. Véase la pág. 689.

4. Aproxímese hasta el término de tercer grado la solución de

a) $\dfrac{dy}{dx} = \dfrac{1}{1-x-y}$, $y(1) = -2$

b) $\ddot{x}(t) + tx(t)\dot{x}(t) + x(t) + t^3 = 0$, $x(0) = 1$, $\dot{x}(0) = 0$

c) $\dot{x} = y + x^2$
$\dot{y} = -x + y$, $x(0) = 0$, $y(0) = 1$

d) $\dot{x}(t) = x(t) + tx(t)$
$\dot{y}(t) = 2tx(t) + y(t)$, $x(0) = a$, $y(0) = b$

e) $\dot{x}(t) = y(t) + t + 1$
$\dot{y}(t) = y(t) + x^2(t)$, $x(0) = 2$, $y(0) = 2$

f) $\ddot{x} + \varepsilon(x^2 - 1)\dot{x} + x = 0$, $x(0) = 2$, $\dot{x}(0) = 0$.

5. *a*) Determínese una solución en serie de

$$(x-1)(x-k)\frac{d^2y}{dx^2} - (x+1-2k)\frac{dy}{dx} + y = 0$$

que satisfaga $y(0) = 1 - 2k$, $y'(0) = 1$.

b) Para $k \neq 0$ la solución en serie obtenida en la parte *a* es *la* solución que satisface la condición inicial. Cuando $k = 0$, pruébese que

$(1-x)^{-1}$ es también una solución que satisface la misma condición inicial.

6. Aproxímese hasta el término de grado 3 la solución de

$$\frac{dy}{dx} = y^2+x^2, \quad y(0) = 1 .$$

6. SOLUCIÓN NUMÉRICA DE LAS ECUACIONES DIFERENCIALES

Hay un gran número de métodos para el cálculo de soluciones de las ecuaciones diferenciales. Presentaremos como ejemplo un método iterativo adecuado para el empleo de calculadoras y que en algunos casos da resultados plenamente satisfactorios. Con las calculadoras modernas de alta velocidad no hay, salvo para algunos problemas especiales, gran dificultad en el cálculo de soluciones particulares. No vamos a efectuar ninguna comparación entre los distintos métodos ni vamos a entrar en los problemas del análisis de error. Es una historia demasiado larga para que comencemos a discutirla aquí. El punto que aquí queremos señalar es que existen métodos para el cálculo de soluciones particulares y que no es ningún problema de mayor cuantía calcular una tabla de valores para una solución particular, cuando la solución del problema ha sido reducida a la de una ecuación diferencial. A menudo se necesitan conocer las propiedades generales de una solución o de un conjunto de soluciones, y aunque una máquina calculadora, cuando se usa inteligentemente, puede ser útil también para tal tarea, puede que no sea suficiente. El uso inteligente de las máquinas y de los métodos numéricos y la investigación de las propiedades generales de soluciones requiere el conocimiento de la teoría de ecuaciones diferenciales.

Consideremos el sistema n-dimensional de ecuaciones diferenciales definido por

6.1 $$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t),$$

y supongamos que queremos calcular la solución de 6.1 que satisface $\mathbf{x}(t_0)=\mathbf{x}^0$ sobre el intervalo $[t_0, T]$. Subdividamos el intervalo en un número finito de partes de longitud h. Sea $t_j = t_0+jh$. Definimos el método por inducción. Supongamos que hemos calculado ya $\mathbf{x}(t_0)$, $\mathbf{x}(t_1)$, ..., $\mathbf{x}(t_n)$, y queremos calcular $\mathbf{x}(t_{n+1}) = \mathbf{x}(t_n+h)$. La función

6.2 $$\varphi(t) = \mathbf{f}(\mathbf{x}(t_n), t_n) + \frac{\mathbf{f}(\mathbf{x}(t_n), t_n)-\mathbf{f}(\mathbf{x}(t_{n-1}), t_{n-1})}{h}(t-t_n)$$

es entonces una aproximación lineal de $\mathbf{f}(\mathbf{x}(t), t)$ que toma los valores

$\mathbf{f}(\mathbf{x}(t_{n-1}), t_{n-1})$ y $\mathbf{f}(\mathbf{x}(t_n), t_n)$ en $t = t_{n-1}$ y $t = t_n$, respectivamente. Usando $\varphi(t)$ para aproximar $\mathbf{f}(\mathbf{x}(t), t)$ sobre el intervalo $[t_n, t_{n+1}]$ y el símbolo "$\cong$" para indicar "es aproximadamente igual a", de 6.1 obtenemos

$$\mathbf{x}(t_{n+1}) - \mathbf{x}(t_n) \cong \int_{t_n}^{t_{n+1}} \varphi(t)\, dt$$
$$= h\mathbf{f}(\mathbf{x}(t_n), t_n) + \tfrac{1}{2}h\left[\mathbf{f}(\mathbf{x}(t_n), t_n) - \mathbf{f}(\mathbf{x}(t_{n-1}), t_{n-1})\right].$$

Usando la notación más conveniente $\mathbf{x}_n = \mathbf{x}(t_n)$, $\mathbf{f}_n = \mathbf{f}(\mathbf{x}_n, t_n)$ y $\Delta \mathbf{f}_n = \mathbf{f}_n - \mathbf{f}_{n-1}$, tenemos entonces

6.3 $$\mathbf{x}_{n+1} \cong \mathbf{x}_n + h\mathbf{f}_n + \tfrac{1}{2}h\Delta\mathbf{f}_n.$$

Nótese que calcular $\mathbf{x}_{n+1}$ requiere que conozcamos $\mathbf{x}_n$ y $\mathbf{x}_{n-1}$. Así, para calcular $\mathbf{x}_2$ debemos conocer $\mathbf{x}_0$ y $\mathbf{x}_1$. El valor de $\mathbf{x}_0$ está dado por la condición inicial, y para calcular $\mathbf{x}_1$ debe utilizarse algún método aproximativo. Un método conveniente es la aproximación por la serie de Taylor discutida en la sección 5. No hemos dado ni intentaremos dar una justificación de este método de cálculo aproximativo de una solución. Parece razonable y podemos esperar para un sistema dado 6.1, que la precisión del método dependerá del tamaño de h y de cuán buena sea en realidad la aproximación obtenida para $\mathbf{x}_1$. También podríamos esperar mejorar el método reemplazando $\varphi(t)$ en 6.2 por un polinomio de segundo grado que tome los valores $\mathbf{f}_n$, $\mathbf{f}_{n-1}$, $\mathbf{f}_{n-2}$ en t_n, t_{n-1}, t_{n-2} (problema 4).

Ilustramos este método (a veces llamado *método de Adam*) aplicándolo primero a una ecuación de primer orden, y luego, a una de segundo.

6.4 Ejemplo. Calcúlese una solución aproximada de

$$\dot{x} = (1-x^2)^{1/2}, \quad x(0) = 0$$

sobre el intervalo [0, 1].

Solución. Tomando $h = 0.100$ y usando los primeros dos términos en la serie de Taylor para la solución, tenemos

$$x_1 = x(h) \cong x(0) + h(1-x^2(0))^{1/2} = h = 0.100.$$

Calculamos entonces los valores paso a paso usando

$$x_{n+1} \cong x_n + hf_n + \tfrac{1}{2}h\Delta f_n.$$

Aquí $f(x) = (1-x^2)^{1/2}$, $t_n = nh$, $x_n = x(t_n)$, $f_n = f(x_n)$, y $\Delta f_n = f_n - f_{n-1}$. La siguiente tabla contiene los resultados de este cálculo con tres cifras significativas. La solución exacta es la función seno, y comparando con la última columna vemos la precisión de nuestros cálculos.

n	t_n	x_n	f_n	Δf_n	sen t_n
0	0.000	0.000	1.000		0.0000
1	0.100	0.100	0.995	−0.005	0.0998
2	0.200	0.199	0.980	−0.015	0.1987
3	0.300	0.296	0.955	−0.025	0.2955
4	0.400	0.389	0.921	−0.034	0.3894
5	0.500	0.479	0.878	−0.043	0.4794
6	0.600	0.565	0.825	−0.053	0.5646
7	0.700	0.645	0.764	−0.061	0.6442
8	0.800	0.718	0.696	−0.068	0.7174
9	0.900	0.784	0.620	−0.076	0.7833
10	1.000	0.842			0.8415

6.5 Ejemplo. Calcúlese una solución aproximada de

$$\ddot{x}+x-x^2 = 0, \qquad x(0) = 0, \qquad \dot{x}(0) = 1$$

sobre el intervalo [0, 0.2].

SOLUCIÓN. Haciendo $\dot{x} = y$ tenemos el siguiente sistema equivalente

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= x^2 - x, \qquad x(0) = 0, \qquad y(0) = 1. \end{aligned}$$

Tomamos $h = 0.05$ y definimos

$$t_n = nh, \qquad x_n = x(t_n), \qquad y_n = y(t_n).$$

Usando los primeros dos términos de la serie de Taylor, obtenemos

$$\begin{aligned} x_1 &= x(h) \cong x(0)+h\dot{x}(0) = h = 0.050 \\ y_1 &= y(h) \cong y(0)+h\dot{y}(0) = y(0) = 1.000. \end{aligned}$$

Por 6.3

$$\begin{aligned} x_{n+1} &\cong x_n+hy_n+\tfrac{1}{2}h\Delta y_n \\ y_{n+1} &\cong y_n+h(x_n^2-x_n)+\tfrac{1}{2}h\Delta(x_n^2-x_n). \end{aligned}$$

De donde obtenemos

n	t_n	x_n	y_n	$x_n^2-x_n$	Δy_n	$\Delta(x_n^2-x_n)$
0	0.000	0.000	1.000	0.000		
1	0.050	0.050	1.000	−0.048	0.000	−0.048
2	0.100	0.100	0.996	−0.090	−0.004	−0.042
3	0.150	0.150	0.990	−0.128	−0.006	−0.038
4	0.200	0.200	0.975			

Problemas

1. Usando $h = 0.05$ en el ejemplo 6.4, calcúlese sen 1 con cuatro cifras significativas.

2. Calcúlese una tabla de valores de la solución de cada una de las siguientes ecuaciones diferenciales usando los valores de h y n que se indican. Siempre que sea posible, compárese con la solución exacta.

a) $\dot{x}(t) = x(t)+2t,\ x(0) = 1,\ h = 0.10,\ n = 5$

b) el mismo problema que (*a*) con $h = 0.05$ y $n = 10$

c) $\dfrac{dy}{dx} = \sqrt{x}+y,\ y(1) = 0,\ h = 0.10,\ n = 5$

d) $\dot{x}(t) = x^2(t)+t,\ x(0) = 1,\ h = 0.10,\ n = 5$

e) $\dot{x} = xy^2$
$\dot{y} = -yx^2,\ x(0) = 1,\ y(0) = 2,\ h = 0.10,\ n = 5.$

3. Calcúlese π por la solución numérica de

$$\frac{dy}{dx} = (1+x^2)^{-1}.$$

4. Supongamos que $\mathbf{x}_n$, $\mathbf{x}_{n-1}$ y $\mathbf{x}_{n-2}$ han sido calculadas. Reemplácese la ecuación $\boldsymbol{\varphi}(t)$ de 6.2 por un polinomio de grado dos que tome los valores $\mathbf{f}_n, \mathbf{f}_{n-1}, \mathbf{f}_{n-2}$ en t_n, t_{n-1}, t_{n-2}. Obténgase de aquí la fórmula

$$\mathbf{x}_{n+1} \cong \mathbf{x}_n + h[\mathbf{f}_n + \tfrac{1}{2}\Delta\mathbf{f}_{n-1} + \tfrac{5}{12}\Delta^2\mathbf{f}_{n-2}]$$

donde $\Delta\mathbf{f}_k = \mathbf{f}_{k+1} - \mathbf{f}_k$ y $\Delta^2\mathbf{f}_k = \Delta(\Delta\mathbf{f}_k) = \Delta\mathbf{f}_{k+1} - \Delta\mathbf{f}_k$. Esta es una fórmula del tipo "paso a paso" para la solución numérica de $\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), t)$ que tiene en cuenta las "segundas diferencias" $\Delta^2\mathbf{f}_k$. Inicialmente se da $\mathbf{x}_0$ y deben entonces calcularse $\mathbf{x}_1$ y $\mathbf{x}_2$ por algún otro método como, por ejemplo, una aproximación por la serie de Taylor.

5. Úsese el método del problema 4 para resolver:

a) Ejemplo 6.4.
b) Problema 2*a*.
c) Problema 2*c*.
d) Problema 2*e*.

7. LOS POLINOMIOS DE LEGENDRE

Muchas funciones especiales han surgido de la física matemática, y en esta sección ilustraremos lo que, para una gran clase de tales funciones, es un estudio típico de algunas de sus propiedades. Los polinomios de

Legendre fueron introducidos en 1784 por Adrien Legendre (1752-1833) y son soluciones de la ecuación diferencial

7.1 $$(1-t^2)\,\ddot{x}(t)-2t\dot{x}(t)+n(n+1)\,x(t) = 0.$$

Esta ecuación diferencial lineal se llama *ecuación diferencial de Legendre.*

Sea n un entero no negativo y busquemos una solución

$$x(t) = \sum_{k=0}^{\infty} a_k t^k,$$

analítica en una vecindad del origen. Sustituyendo las series en 7.1, obtenemos

$$\sum_{k=0}^{\infty} \{(1-t^2)\,k(k-1)a_k t^{k-2}-2tka_k t^{k-1}+n(n+1)a_k t^k\} = 0.$$

Al igualar el coeficiente de t^k a cero, obtenemos

7.2 $$(k+1)(k+2)a_{k+2} = [k(k+1)-n(n+1)]a_k.$$

Las soluciones principales, que denotamos por u_n y v_n, corresponden a las condiciones iniciales $(u_n(0), \dot{u}_n(0)) = (1, 0)$ y $(v_n(0), \dot{v}_n(0)) = (0, 1)$. Examinando las series para u_n, notamos en primer lugar que la serie es par: $a_1 = 0$ y se sigue de 7.2 que todos los coeficientes impares son cero. La ecuación para los coeficientes pares es

7.3 $$(2k+1)(2k+2)a_{2k+2} = [2k(2k+1)-n(n+1)]a_{2k}.$$

Si n es un entero par, la serie termina con $a_n t^n$ $(a_{n+2} = 0)$ y u_n es un polinomio par de grado n. Si n es impar, se sigue de 7.3 y el criterio de la razón, que la serie converge en el intervalo abierto $\langle -1, 1\rangle$ (problema 1*a*). Se puede probar que diverge en los puntos extremos ± 1 (problema 1*b*). Análogamente, para v_n, la serie es impar, es un polinomio impar de grado n si n es impar, y en otro caso es una serie infinita cuyo intervalo de convergencia es $\langle -1, 1\rangle$.

Restringiremos nuestra atención a estas soluciones polinomiales que denotaremos por y_n. Entonces

$$y_n(t) = b_0 t^n + b_1 t^{n-2} + \cdots = \sum_{k=0}^{[\frac{1}{2}n]} b_k t^{n-2k},$$

donde el máximo entero $[\frac{1}{2}n]$ es igual a $\frac{1}{2}n$ si n es par y a $\frac{1}{2}(n-1)$ si n es impar. Sustituyendo en 7.1 obtenemos la fórmula de recursión

$$b_k = -\frac{(n-2k+2)(n-2k+1)}{2k(2n-2k+1)}\,b_{k-1}.$$

De donde las soluciones polinomiales son

$$y_n(t) = b_0\left[t^n - \frac{n(n-1)}{2(2n-1)}t^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2\cdot 4(2n-1)(2n-3)}t^{n-4} - \cdots\right]$$

$$= \frac{(n!)^2}{(2n)!}b_0 \sum_{k=0}^{[\frac{1}{2}n]} (-1)^k \frac{(2n-2k)!}{k!(n-k)!(n-2k)!}t^{n-2k}.$$

El *polinomio de Legendre* P_n de grado n se define por

7.4
$$P_n(t) = \frac{1}{2^n}\sum_{k=0}^{[\frac{1}{2}n]} (-1)^k \frac{(2n-2k)!}{k!(n-k)!(n-2k)!}t^{n-2k}$$

que corresponde a tomar $b_0 = \dfrac{(2n)!}{2^n(n!)^2}$. Esta elección de b_0 es conveniente para muchos propósitos (véase, por ejemplo, 7.5 y problema 4). He aquí unos pocos de los primeros polinomios de Legendre:

$$P_0(t) = 1, \quad P_1(t) = t, \quad P_2(t) = \tfrac{1}{2}(3t^2-1),$$

$$P_3(t) = \tfrac{1}{2}(5t^3-3t), \quad P_4(t) = \tfrac{1}{8}(35t^4-30t^2+3)$$

$$P_5(t) = \tfrac{1}{8}(63t^5-70t^3+15t).$$

Una consecuencia (problema 5) de esta elección de b_0 es que

7.5
$$P_n(1) = 1, \quad P_n(-1) = (-1)^n$$

$$P_{2n+1}(0) = 0, \quad P_{2n}(0) = (-1)^n \frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n} = (-1)^n\frac{(2n-1)!!}{(2n)!!}.$$

Lo que ahora queremos hacer es ilustrar de qué modo podemos derivar de la propia ecuación diferencial algunas propiedades generales de los polinomios de Legendre.

Podemos escribir la ecuación diferencial de Legendre 7.1 en la forma

$$\frac{d}{dt}[(1-t^2)\dot{x}(t)] + n(n+1)x(t) = 0.$$

Si definimos $L(x)(t) = \dfrac{d}{dt}((1-t^2)\dot{x}(t))$ y $\lambda = -n(n+1)$, podemos expresar esto por

7.6
$$L(x) = \lambda x,$$

lo que se llama un "problema de valor característico". Los polinomios de Legendre son soluciones de la ecuación de Legendre; es decir

7.7 $$L(P_n) = -n(n+1)P_n.$$

Así pues, por analogía con el problema del valor característico para matrices $-n(n+1)$ se llama valor característico y P_n se llama una función característica. Es ésta una analogía provechosa que ha sido explotada y nos proporciona una vía de acceso general al estudio de una amplia clase de ecuaciones diferenciales lineales y de problemas de frontera de interés en la física matemática y en otras muchas ramas (sistemas de Sturm-Liouville). Nuestro objetivo es el mucho más modesto: queremos ilustrar simplemente un caso especial. Tenemos la siguiente identidad para el operador o transformación L: sean x_1 y x_2 un par de funciones diferenciables sobre $\langle -\infty, \infty\rangle$. Entonces

7.8 $$x_1 L(x_2) - x_2 L(x_1) = \frac{d}{dt}[(1-t^2)(x_1 \dot{x}_2 - x_2 \dot{x}_1)].$$

De esta simple identidad, llamada *identidad de Lagrange*, podemos deducir la siguiente importante propiedad de los polinomios de Legendre:

7.9 $$\int_{-1}^{1} P_n P_m = 0 \quad \text{para} \quad n \neq m.$$

Por 7.7 y 7.8 tenemos

$$[m(m+1) - n(n+1)] \int_{-1}^{1} P_n P_m = \int_{-1}^{1} [P_m L(P_n) - P_n L(P_m)]$$

$$= \int_{-1}^{1} \frac{d}{dt}[(1-t^2)(P_m(t)\dot{P}_n(t) - P_n(t)\dot{P}_m(t))]\,dt = 0.$$

Bajo la hipótesis $n \neq m$ (n y m son enteros no negativos), obtenemos 7.9. Esta propiedad 7.9 se llama "ortogonalidad" de los polinomios de Legendre sobre el intervalo $[-1, 1]$. En la sección 9 veremos más sobre la significación de esta propiedad.

Podemos usar ahora la ortogonalidad 7.9 y el teorema de existencia y unicidad para ecuaciones diferenciales para establecer una propiedad de las raíces de P_n.

7.10 Teorema. *Todas las raíces de P_n son distintas y se encuentran en el intervalo $\langle -1, 1\rangle$.*

PRUEBA. Sea R_m un polinomio de grado m. Elíjase c_m de manera que el coeficiente de t^m en $c_m P_m(t)$ sea el mismo que el que tiene en R_m. Elíjase

después c_{m-1} de modo que el coeficiente de t^{m-1} en $c_{m-1}P_{m-1}(t)$ sea el mismo que el que tiene en $R_m(t)-c_m P_m(t)$. En esta forma vemos que podemos determinar $c_m, \ldots, c_0$ de modo que

7.11 $$R_m = c_0 P_0 + c_1 P_1 + \ldots + c_m P_m .$$

Todo polinomio de grado m puede expresarse como una combinación lineal de los polinomios de Legendre $P_0, \ldots, P_m$. Se sigue entonces por 7.9 que, si R_m es un polinomio cualquiera de grado m

7.12 $$\int_{-1}^{1} R_m P_n = 0 \quad \text{para} \quad m < n .$$

Estamos ahora preparados para investigar las raíces de P_n. Como P_n es una solución de la ecuación diferencial 7.1 no puede tener una raíz múltiple en $\langle -1, 1\rangle$. Si tuviera una raíz múltiple r en $\langle -1, 1\rangle$, entonces $P_n(r) = P_n'(r) = 0$. Pero la única solución de 7.1 sobre $\langle -1, 1\rangle$ que satisface esta condición inicial es la función cero. Luego las raíces de P_n en $\langle -1, 1\rangle$ son distintas. Supongamos que las raíces en $\langle -1, 1\rangle$ son $t_1, t_2, \ldots, t_m$. Definamos

$$R_m(t) = (t-t_1)(t-t_2)\cdots(t-t_m).$$

R_m tiene las mismas raíces en el intervalo $\langle -1, 1\rangle$ que P_n y tiene el mismo signo que P_n en $t = 1$ (ambos son positivos). Luego

$$\int_{-1}^{1} R_m P_n > 0 .$$

Como $m \leqslant n$, se sigue de 7.12 que $m = n$, lo que completa la prueba. Para otra prueba del teorema 7.10 véase problema 7.

Problemas

1. Consideremos la serie $\sum_{k=0}^{\infty} a_{2k} t^{2k}$ donde $a_0 = 1$ y los coeficientes a_{2k} están definidos por 7.3. Pruébese que:

a) La serie converge en el intervalo $\langle -1, 1\rangle$.

b) La serie diverge en $t = \pm 1$.

2. Llénense los detalles faltantes de la discusión que condujo a la fórmula para las soluciones polinomias de la ecuación de Legendre.

3. Derívese una expansión en serie para la solución general de la ecuación de Legendre sobre el intervalo $\langle -1, 1\rangle$ suponiendo solamente que n es un número real.

4. Pruébese que

$$(1-2tx+x^2)^{-1/2} = P_0(t) + P_1(t)x + \cdots = \sum_{n=0}^{\infty} P_n(t)x^n$$

para cualquier valor de t y para $|x|$ suficientemente pequeña. ($(1-2tx+x^2)^{-\frac{1}{2}}$ se llama *función generadora* de los polinomios de Legendre.)

5. Pruébese 7.5. *Sugerencia:* úsese la función generadora del problema 4.

6. Pruébese que

$$P_n(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}(t^2-1)^n.$$

A esto se llama *fórmula de Rodrigues* (1815).

7. Pruébese el teorema 7.10 usando la fórmula de Rodrigues del problema 6 y el teorema de Rolle.

8. Pruébese que:

$$\int_{-1}^{1} P_n{}^2 = \frac{2}{2n+1}, \qquad n = 0, 1, \cdots$$

Sugerencia: puede establecerse esto usando la fórmula de Rodrigues y la integración por partes. Una prueba más ingeniosa puede basarse en el cuadrado de la expansión en problema 4 y la propiedad de ortogonalidad de los polinomios de Legendre.

9. Pruébese que:

a) $(n+1)P_{n+1}(t)-(2n+1)tP_n(t)+nP_{n-1}(t)=0$

b) $tP_n{}'(t)-P'_{n-1}(t) = nP_n(t)$

c) $P'_{n+1}(t)-tP_n{}'(t) = (n+1)P_n(t)$.

Sugerencia: *a*) considérese $(1-2tx+x^2)\dfrac{\partial V}{\partial x} = (x-t)V$ donde $V = (1-2tx+x^2)^{-1/2}$.

b) Considérese $x\dfrac{\partial V}{\partial x} = (t-x)\dfrac{\partial V}{\partial t}$.

10. Pruébese que:

$$c_0P_0+c_1P_1+\cdots+c_nP_n = 0$$

implica $c_0=c_1=\cdots=c_n=0$ (los polinomios de Legendre son linealmente independientes).

11. Si f es un polinomio de grado n, pruébese que

$$f = c_0P_0+c_1P_1+\cdots+c_nP_n$$

donde $c_k = \dfrac{2k+1}{2}\displaystyle\int_{-1}^{1} f\,P_k$.

12. Pruébese que:

a) Las funciones $\varphi_n = e^{int}$, $n = 0, \pm 1, \ldots$, son ortogonales sobre el intervalo $[-\pi, \pi]$.

b) $c_{-m}\varphi_{-m} + c_{-m+1}\varphi_{-m+1} \cdots + c_0\varphi_0 + c_1\varphi_1 + \cdots + c_n\varphi_n = 0$, implica $c_{-m} = c_{-m+1} = \cdots = c_0 = c_1 = \cdots = c_n = 0$ (las funciones φ_k, $k = -m, \ldots, n$, son linealmente independientes).

c) sen t, sen $2t, \ldots$, son ortogonales y linealmente independientes sobre el intervalo $[0, \pi]$.

13. *La ecuación diferencial de Hermite es*

$$\ddot{x}(t) - 2t\dot{x}(t) + 2nx(t) = 0$$

La solución polinomial H_n de esta ecuación diferencial en que el coeficiente de t^n es 2^n se llama *polinomio de Hermite de grado n.* Pruébese que:

a) $$H_n(t) = \sum_{k=0}^{[\frac{1}{2}n]} (-1)^k \frac{n!}{k!(n-2k)!} (2t)^{n-2k}$$

b) $$e^{-x^2+2tx} = \sum_{n=0}^{\infty} H_n(t) \frac{x^n}{n!}$$

c) $$\int_{-\infty}^{\infty} e^{-t^2} H_n(t) H_m(t)\, dt = 0, \quad n \neq m.$$

Sugerencia: nótese que $L(H_n(t)) = -2ne^{-t^2}H_n(t)$ donde $L(H_n(t)) = \frac{d}{dt}(e^{-t^2}H_n'(t))$.

8. SERIES DE FOURIER

Las series de Fourier tuvieron su origen en el siglo XVIII en el trabajo de Daniel Bernoulli (1700-1782) en conexión con el problema de una cuerda vibrante. Su trabajo (*Sur les Cordes Vibrantes*, 1753) planteó la cuestión de si toda función f que es continua sobre un intervalo $[0, \pi]$ y se anula en los puntos extremos puede representarse por una serie de senos; es decir, ¿puede uno encontrar constantes b_n tales que

8.1 $$f(x) = \sum_{n=1}^{\infty} b_n \text{ sen } nx$$

sobre $[0, \pi]$? Basándose en la intuición física, Bernoulli dijo que sí. Los matemáticos más distinguidos de entonces dijeron que no. Sus argumentos contra Bernoulli se basaban en una intuición matemática ingenua y en

un concepto limitado de lo que es una función. En 1822, Juan Bautista Fourier (1768-1830) usó series trigonométricas en su teoría de la conducción del calor (*La Théorie Analytique de la Chaleur*). Su trabajo hizo surgir el problema de si las funciones f continuas sobre $[-\pi, \pi]$ con $f(-\pi) = f(\pi)$ pueden representarse por series de senos y cosenos; es decir, ¿pueden encontrarse constantes a_n, b_n tales que

8.2 $$f(x) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx)$$

sobre $[-\pi, \pi]$? El uso de Fourier de estas series era puramente formal, y ello renovó la controversia que había comenzado con Bernoulli. Esta controversia tuvo un profundo efecto sobre las matemáticas y condujo a la clarificación de muchas cuestiones relacionadas con las series y las funciones.

Es mucho más conveniente expresar el problema de representación que plantea 8.2 en términos de la exponencial compleja. Supongamos que f es continua sobre $[-\pi, \pi]$ y que $f(-\pi) = f(\pi)$. Nos preguntamos entonces si podemos encontrar coeficientes c_n con la propiedad de que

8.3 $$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

8.4 Teorema. *Si la serie* $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ *converge uniformemente a* f *sobre* $[-\pi, \pi]$, *entonces los coeficientes* c_n *vienen dados por*

8.5 $$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx.$$

PRUEBA. Como la serie converge uniformemente a f sobre $[-\pi, \pi]$ podemos integrar término a término. Usando la propiedad de ortogonalidad

$$\int_{-\pi}^{\pi} e^{inx} e^{-imx}\, dx = \begin{cases} 0, & m \neq n \\ 2\pi, & m = n \end{cases}$$

para calcular los coeficientes, obtenemos

$$\int_{-\pi}^{\pi} f(x) e^{-imx}\, dx = \sum_{n=-\infty}^{\infty} c_n \int_{-\pi}^{\pi} e^{inx} e^{-imx}\, dx = 2\pi c_m.$$

De donde

$$c_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-imx}\, dx.$$

Llamamos a éstos *coeficientes de Fourier* de f.

Si preferimos trabajar con la serie trigonométrica 8.2, tenemos

$$\sum_{n=-\infty}^{\infty} c_n e^{inx} = \sum_{n=-\infty}^{\infty} c_n(\cos nx + i \operatorname{sen} nx)$$

$$= c_0 + \sum_{n=1}^{\infty} [(c_n + c_{-n}) \cos nx + i(c_n - c_{-n}) \operatorname{sen} nx].$$

Por tanto

8.6

$$a_n = c_n + c_{-n} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx \qquad n = 0, 1, 2, \cdots$$

$$b_n = i(c_n - c_{-n}) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx\, dx \qquad n = 1, 2, \cdots$$

Estos son los *coeficientes de Fourier* de f para la serie trigonométrica 8.2.

8.7 Corolario. *Si la serie* $\frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx)$ *converge uniformemente a* f *sobre* $[-\pi, \pi]$, *entonces los coeficientes* a_n *y* b_n *están dados por* 8.6.

Si f es continua sobre $[-\pi, \pi]$, entonces los coeficientes de Fourier c_n, a_n y b_n están definidos y la serie correspondiente se llama *serie de Fourier de* f. Escribimos entonces

$$f \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

o bien

$$f \sim \tfrac{1}{2}a_0 + \sum_{n=0}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx),$$

donde c_n, a_n y b_n son los coeficientes de Fourier de f dados por 8.5 y 8.6. Esto es simplemente una asociación formal de una serie con una función y nada queda implicado respecto a la convergencia de la serie. Uno de los problemas en las series de Fourier es el de procurar una interpretación de " $\sim$ ". Cuándo, por ejemplo, podemos reemplazar " $\sim$ " por " $=$ " donde " $=$ " significa convergencia uniforme a f o algún tipo más débil de convergencia. Se sabe mucho acerca de este problema, pero no podemos entrar aquí en muchos detalles. Hay muchas referencias excelentes (unas cuantas de ellas aparecen en la bibliografía en la página 777) y sólo queremos presentar una breve introducción a este importante tema.

La primera pregunta que contestaremos concierne a la unicidad de la serie de Fourier de f. Supongamos que f y g son un par de funciones

continuas sobre $[-\pi, \pi]$. Queremos probar que las series de Fourier son las mismas si y sólo si $f = g$. Es inmediato que si $f = g$ entonces f y g tienen los mismos coeficientes de Fourier. El problema recíproco es algo más difícil.

8.8 Teorema. *Si f y g son continuas sobre $[-\pi, \pi]$ y tienen los mismos coeficientes de Fourier, entonces $f = g$ sobre $[-\pi, \pi]$.*

PRUEBA. Se sigue de inmediato el enunciado si demostramos que la sola función continua con coeficientes de Fourier cero es la función cero. Entonces, si los coeficientes de Fourier de f y g son iguales, los coeficientes de Fourier de $f-g$ son todos cero. Luego $f-g=0$ y $f=g$.

Supongamos que hay una función continua h distinta de cero y continua sobre $[-\pi, \pi]$ con la propiedad de que

$$\int_{-\pi}^{\pi} h(x)e^{imx}\,dx = 0, \qquad m = 0, \pm 1, \pm 2, \cdots$$

Claramente, la conjugada compleja de h también tiene esta propiedad de modo que podemos suponer que hay una función real distinta de cero h con esta propiedad. Extendamos la definición de h a $\langle -\infty, \infty\rangle$ haciendo h periódica de periodo 2π; es decir, para cada x en $[-\pi, \pi]$ y todo entero k se define $h(x+2k\pi) = h(x)$. Entonces para cualquier c (problema 1)

8.9
$$\int_{c-\pi}^{c+\pi} h(x)e^{imx}\,dx = \int_{-\pi}^{\pi} h(x)e^{imx}\,dx = 0$$

para toda $m = 0, \pm 1, \pm 2, \ldots$. Como $h \neq 0$, entonces, para algún c, $h(c) \neq 0$. (Como h se supuso continua sobre $[-\pi, \pi]$, podemos suponer que c está en $\langle -\pi, \pi\rangle$, y la función extendida h es por tanto continua en c.) Por 8.9 la propiedad de tener coeficientes de Fourier cero es invariante bajo las traslaciones y, además, no cambia por la multiplicación por -1. Podemos, por tanto, suponer que $h(0) > 0$ y que h es continua en 0. Entonces, para algún $\delta > 0$, $h(x) > \frac{1}{2}h(0) > 0$ cuando $|x| < \delta$. Podemos suponer $0 < \delta < \pi$. Definamos

$$P_n(x) = (1 - \cos\delta + \tfrac{1}{2}(e^{ix} + e^{-ix}))^n.$$

No es difícil probar (problema 2) que $1 - \cos\delta + \frac{1}{2}(e^{ix} + e^{-ix}) = 1 - \cos\delta + \cos x$ es mayor que 1 cuando $|x| < \delta < \pi$ y es menor que 1 en valor absoluto cuando $\delta < |x| \leqslant \pi$. Como todos los coeficientes de Fourier de h son cero,

$$\int_{-\pi}^{\pi} h(x)P_n(x)\,dx = 0 \quad \text{para todo } n = 1, 2, \cdots$$

Pero

$$\int_{-\pi}^{\pi} h(x)P_n(x)\,dx$$

$$= \int_{-\delta}^{\delta} h(x)P_n(x)\,dx + \int_{-\pi}^{-\delta} h(x)P_n(x)\,dx + \int_{\delta}^{\pi} h(x)P_n(x)\,dx.$$

Como (problema 3) $\int_{-\delta}^{\delta} h(x)P_n(x)\,dx \to \infty$ cuando $n \to \infty$ e

$$\left|\int_{-\pi}^{-\delta} h(x)P_n(x)\,dx + \int_{\delta}^{\pi} h(x)P_n(x)\,dx\right| < \int_{-\pi}^{\pi} |h(x)|\,dx \quad \text{para todo } n,$$

$\int_{-\pi}^{\pi} h(x)P_n(x)\,dx \to \infty$ cuando $n \to \infty$. Esto es una contradicción y, por tanto, $h = 0$ es la única función continua sobre $[-\pi, \pi]$ con la propiedad de que todos sus coeficientes de Fourier son cero. Y esto completa la prueba.

El siguiente corolario nos demuestra la importancia del anterior teorema.

8.10 Corolario. *Si f es continua sobre $[-\pi, \pi]$ y si su serie de Fourier converge uniformemente sobre $[-\pi, \pi]$, entonces converge a f.*

PRUEBA. Sea $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ la serie de Fourier de f, y supongamos que $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ converge uniformemente sobre $[-\pi, \pi]$ a g. Entonces $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ es también (teorema 8.4) la serie de Fourier de g. Por tanto, por el teorema 8.8 $f = g$, lo que completa la prueba.

Una función f sobre cualquier intervalo $\mathscr{I}$ se dice que es *lisa a trozos* sobre $\mathscr{I}$ si tanto f como f' son continuas a trozos sobre $\mathscr{I}$. Para muchas aplicaciones el siguiente resultado es un teorema de representación adecuado.

8.11 Teorema. *Supongamos que f es lisa a trozos sobre $[-\pi, \pi]$ y luego es extendida a $\langle -\infty, \infty \rangle$ de forma que es periódica de periodo 2π. Entonces la serie de Fourier de f converge en cada punto x a $\frac{1}{2}(f(x-0)+f(x+0))$.*[1]

PRUEBA. Sea

$$s_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}$$

[1] $f(x+0)$ es el límite a la derecha de f en x: $f(x+0) = \lim_{t \to 0^+} f(x+t)$, y $f(x-0)$ es el límite izquierdo de f en x.

la n-ésima suma parcial de la serie de Fourier de f. Entonces por los problemas 5 y 1

$$s_n(x) = \sum_{k=-n}^{n} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u) e^{-iku} e^{ikx}\, du$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u) \sum_{k=-n}^{n} e^{-ik(u-x)}\, du$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u) \frac{\operatorname{sen}\,(n+\frac{1}{2})\,(u-x)}{\operatorname{sen}\,\frac{1}{2}(u-x)}\, du$$

$$= \frac{1}{2\pi} \int_{-\pi-x}^{\pi-x} f(t+x) \frac{\operatorname{sen}\,(n+\frac{1}{2})t}{\operatorname{sen}\,\frac{1}{2}t}\, dt$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t+x) \frac{\operatorname{sen}\,(n+\frac{1}{2})t}{\operatorname{sen}\,\frac{1}{2}t}\, dt.$$

De donde obtenemos

8.12 $$s_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t+x) \frac{\operatorname{sen}\,(n+\frac{1}{2})t}{\operatorname{sen}\,\frac{1}{2}t}\, dt.$$

Definamos

$$h(t) = \frac{f(x+t)-f(x+0)}{2 \operatorname{sen} \frac{1}{2}t}.$$

Es claro que h es continua a trozos (en realidad, lisa a trozos) sobre $\langle 0, \pi]$. Necesitamos saber que h es continua a trozos sobre $[0, \pi]$. Ahora bien

$$\lim_{t\to 0^+} h(t) = \lim_{t\to 0^+} \frac{f(x+t)-f(x+0)}{t} \frac{t}{2 \operatorname{sen} \frac{1}{2}t}$$

$$= \lim_{t\to 0^+} \frac{f(x+t)-f(x+0)}{t}$$

si este límite existe. Pero según el teorema del valor medio y para t suficientemente pequeña

$$\frac{f(x+t)-f(x+0)}{t} = f'(x+\theta_t)$$

para algún $0<\theta_t<t$. Como se supuso que f era lisa a trozos, se sigue entonces que

$$\lim_{t\to 0^+} h(t) = f'(x+0).$$

Este resultado puede también obtenerse directamente usando la regla de l'Hopital (pág. 128). Por tanto, h es continua a trozos en $[0, \pi]$ y, por el problema $4b$,

$$\lim_{n\to\infty} \frac{1}{\pi}\int_0^{\pi} h(t)\operatorname{sen}(n+\tfrac{1}{2})t\,dt = \lim_{n\to\infty} \frac{1}{2\pi}\int_0^{\pi} f(x+t)\frac{\operatorname{sen}(n+\frac{1}{2})t}{\operatorname{sen}\frac{1}{2}t}\,dt$$

$$- \lim_{n\to\infty} \frac{1}{2\pi}\int_0^{\pi} f(x+0)\frac{\operatorname{sen}(n+\frac{1}{2})t}{\operatorname{sen}\frac{1}{2}t}\,dt = 0\,.$$

Se sigue entonces del problema 6 que

$$\lim_{n\to\infty} \frac{1}{2\pi}\int_0^{\pi} f(x+t)\frac{\operatorname{sen}(n+\frac{1}{2})t}{\operatorname{sen}\frac{1}{2}t}\,dt = \tfrac{1}{2}f(x+0)\,.$$

Análogamente,

$$\lim_{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{0} f(x+t)\frac{\operatorname{sen}(n+\frac{1}{2})t}{\operatorname{sen}\frac{1}{2}t}\,dt = \tfrac{1}{2}f(x-0).$$

Combinando estos dos límites, tenemos (véase 8.12)

$$\lim_{n\to\infty} s_n(x) = \tfrac{1}{2}\{f(x-0)+f(x+0)\},$$

lo que completa la prueba.

Al discutir la serie de Fourier de una función tomábamos $[-\pi, \pi]$ como intervalo básico. Lo mismo podíamos haber tomado cualquier intervalo cerrado finito $-\frac{1}{2}l, \frac{1}{2}l$. Cambiando a este intervalo tenemos

8.13 $$f \sim \sum_{n=-\infty}^{\infty} c_n e^{i2n\pi(x/l)}$$

donde

81.4 $$c_n = \frac{1}{l}\int_{-l/2}^{l/2} f(x)e^{-i2n\pi(x/l)}\,dx\,;$$

o bien

8.15 $$f \sim \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty}\left(a_n \cos 2n\pi\frac{x}{l} + b_n \operatorname{sen} 2n\pi\frac{x}{l}\right)$$

donde

8.16

$$a_n = \frac{2}{l}\int_{-l/2}^{l/2} f(x)\cos 2n\pi\frac{x}{l}\,dx$$

$$b_n = \frac{2}{l}\int_{-l/2}^{l/2} f(x)\,\mathrm{sen}\, 2n\pi\frac{x}{l}\,dx.$$

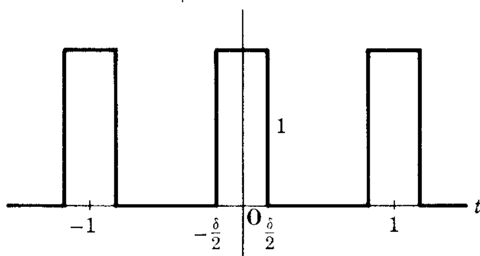


FIGURA 1

8.17 Ejemplo. Sea f la serie periódica de pulsos rectangulares unitarios que mostramos en la figura 1. Por conveniencia tomamos como periodo la unidad de tiempo. Calcúlese la serie de Fourier de f.

SOLUCIÓN. Los coeficientes de Fourier de f son

$$c_n = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(t)e^{-i2n\pi t}\,dt$$

$$= \int_{-\frac{1}{2}\delta}^{\frac{1}{2}\delta} e^{-i2n\pi t}\,dt = \frac{1}{n\pi}\,\mathrm{sen}\, n\pi\delta, \quad n \neq 0$$

y

$$c_0 = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(t)\,dt = \delta.$$

De donde

$$f(t) \sim \delta + \sum_{n=-\infty}^{-1} \frac{\mathrm{sen}\, n\pi\delta}{n\pi} e^{i2n\pi t} + \sum_{n=1}^{\infty} \frac{\mathrm{sen}\, n\pi\delta}{n\pi} e^{i2n\pi t}$$

$$= \delta + 2\delta \sum_{n=1}^{\infty} \frac{\mathrm{sen}\, n\pi\delta}{n\pi\delta}\cos 2n\pi t.$$

En $t = \pm\frac{1}{2}\delta$ la serie de Fourier converge al valor medio $\frac{1}{2}$. Las amplitudes relativas de las armónicas $2n\pi t$ varían como $\frac{1}{x}\,\mathrm{sen}\, x$, donde $x = n\pi\delta$.

Por tanto, podemos esperar que cuanto más pequeño sea δ mayor será la importancia de las armónicas más altas. Esto se sabe muy bien por los diseñadores de amplificadores de pulsos rectangulares.

Mencionamos al principio que a menudo se está interesado en la representación de una función f sobre un intervalo [0, 1] por una serie de senos. Si definimos $f(-x) = -f(x)$, entonces obtenemos esto como un caso particular de las más generales series trigonométricas de Fourier. Sobre el intervalo $[-1, 1]$

$$f(x) \sim \frac{1}{2}a_0 + \sum_{n=1}^{\infty}\left(a_n \cos n\pi\frac{x}{l} + b_n \operatorname{sen} n\pi\frac{x}{l}\right).$$

Pero como f es una función impar, $a_n = 0$ para todo n (problema $8a$) y

$$f(x) \sim \sum_{n=1}^{\infty} b_n \operatorname{sen} n\pi\frac{x}{l}$$

donde

$$b_n = \frac{2}{l}\int_0^l f(x) \operatorname{sen} n\pi\frac{x}{l}\,dx.$$

Se sigue entonces del teorema 8.11 que si f es lisa a trozos sobre [0, 1] esta serie de senos converge para cada x a $\frac{1}{2}(f(x-0)+f(x+0))$.

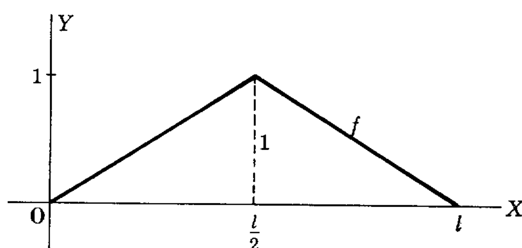


FIGURA 2

8.18 Ejemplo. Represéntese la función f que aparece en la figura 2 por una serie de senos sobre el intervalo [0, 1].

SOLUCIÓN. Como $f(l-x) = f(x)$, tenemos que

$$b_n = \frac{2}{l}\int_0^l f(x) \operatorname{sen} n\pi\frac{x}{l}\,dx$$

$$= \frac{2}{l}\int_0^{l/2} f(x)\,\text{sen}\,n\pi\frac{x}{l}\,dx + \frac{2}{l}\int_0^{l/2} f(x)\,\text{sen}\,n\pi\left(1-\frac{x}{l}\right)dx$$

$$= \frac{2}{l}\,[1-(-1)^n]\int_0^{l/2} f(x)\,\text{sen}\,n\pi\frac{x}{l}\,dx\,.$$

De donde $b_{2n} = 0$ y

$$b_{2n+1} = \frac{4}{l}\int_0^{l/2}\frac{2}{l}x\,\text{sen}\,(2n+1)\pi\frac{x}{l}\,dx = (-1)^n\frac{8}{(2n+1)^2\pi^2}\,.$$

Por tanto

$$f(x) = \frac{8}{\pi^2}\left[\text{sen}\,\frac{\pi x}{l} - \frac{1}{9}\,\text{sen}\,\frac{3\pi x}{l} + \frac{1}{25}\,\text{sen}\,\frac{5\pi x}{l} + \cdots\right]$$

$$= \frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^2}\,\text{sen}\,(2n+1)\frac{\pi x}{l}\,.$$

Como el término general en la serie es en valor absoluto menor que $1/(2n+1)^2$, sabemos que la serie converge uniformemente sobre $[0, 1]$. Por tanto, por el corolario 8.10 la serie converge uniformemente a f sobre $[0, 1]$. En particular, tenemos al evaluar la serie en $x = \frac{1}{2}$ que

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots$$

Problemas

1. Si f es continua y periódica de periodo T pruébese que para todo c

$$\int_0^T f = \int_c^{c+T} f\,.$$

2. Pruébese que $1 - \cos\delta + \cos x > 1$ cuando $|x| < \delta < \pi$ y

$$|1 - \cos\delta + \cos x| < 1$$

cuando $0 < \delta < |x| \leqslant \pi$.

3. Con $h(x)$, $P_n(x)$ y δ definidas como en la prueba del teorema 8.8 pruébese que

a) $\displaystyle\int_{-\delta}^{\delta} h(x)\,P_n(x)\,dx \to \infty$ cuando $n \to \infty$

b) $\displaystyle\left|\int_{-\pi}^{-\delta} h(x)\,P_n(x)\,dx + \int_{\delta}^{\pi} h(x)\,P_n(x)\,dx\right| < \int_{-\pi}^{\pi} |h(x)|\,dx\,.$

4. *a*) Si f es continua sobre $[a, b]$ pruébese que

$$\lim_{w\to\infty}\int_a^b f(t)\operatorname{sen} wt\,dt = 0.$$

Sugerencia: pruébese primero que

$$2\int_a^b f(t)\operatorname{sen} wt\,dt = -\int_{a-\varepsilon}^{a} f(t+\varepsilon)\operatorname{sen} wt\,dt$$
$$+\int_a^{b-\varepsilon}[f(t)-f(t+\varepsilon)]\operatorname{sen} wt\,dt + \int_{b-\varepsilon}^{b} f(t)\operatorname{sen} wt\,dt,$$

donde $\varepsilon = \dfrac{\pi}{w}$.

b) Conclúyase que la parte *a*) se verifica si f es continua a trozos sobre $[a, b]$.

5. Pruébese que $\displaystyle\sum_{k=-n}^{n} e^{-ik\theta} = \frac{\operatorname{sen}(n+\frac{1}{2})\theta}{\operatorname{sen}\frac{1}{2}\theta}$.

6. Pruébese que

$$\int_0^\pi \frac{\operatorname{sen}(n+\frac{1}{2})t}{\operatorname{sen}\frac{1}{2}t}\,dt = \pi.$$

7. Derívense

a) 8.14 *b*) 8.16.

8. Si f es integrable sobre $[-\pi, \pi]$ pruébese que:

a) Si f es impar, entonces $f(x) \sim \displaystyle\sum_{n=1}^{\infty} b_n \operatorname{sen} nx$, donde

$$b_n = \frac{2}{\pi}\int_0^\pi f(x)\operatorname{sen} nx\,dx$$

b) Si f es par, $f(x) \sim \frac{1}{2}a_0 + \displaystyle\sum_{n=1}^{\infty} a_n \cos nx$, donde

$$a_n = \frac{2}{\pi}\int_0^\pi f(x)\cos nx\,dx.$$

9. Encuéntrense las series de Fourier de

a) $f(x) = x^2$ sobre $[-\pi, \pi]$ *b*) $f(x) = x^2$ sobre $\left[\dfrac{-l}{2}, \dfrac{1}{2}\right]$

c) $f(x) = e^x$ sobre $[-1, 1]$ *d*) $f(x) = x$ sobre $[-\pi, \pi]$
e) $f(x) = |x|$ sobre $[-\pi, \pi]$
f) $f = -1$ sobre $[-\pi, 0]$, $f = 1$ sobre $\langle 0, \pi]$
g) $f = 0$ sobre $[-1, 0]$, $f(x) = x$ para x en $[0, 1]$.

10. Discútase la representación en serie de cosenos de una función f sobre el intervalo $[0, l]$. Pruébese para

$$f(x) \sim \frac{1}{2} a_0 + \sum_{n=1}^{\infty} a_n \cos n\pi \frac{x}{l}$$

que

$$a_n = \frac{2}{l} \int_0^l f(x) \cos n\pi \frac{x}{l} \, dx .$$

11. Proporciónense las representaciones en serie de senos y las representaciones en serie de cosenos de cada una de las siguientes funciones:

a) $f = 1$, sobre $[0, \pi]$ *b*) $f = I$ sobre $[0, \pi]$
c) $f =$ sen sobre $[0, \pi]$ *d*) $f =$ cos sobre $[0, \pi]$
e) $f(x) = x(\pi - x)$ sobre $[0, \pi]$.

12. Si f y g son reales y continuas sobre $[a, b]$ pruébese que

$$\left(\int_a^b fg\right)^2 \leqslant \left(\int_a^b f^2\right)\left(\int_a^b g^2\right).$$

13. Suponiendo que f es real y continua sobre $[-\pi, \pi]$ establézcase la *desigualdad de Bessel*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f^2 \geqslant \sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{1}{2} \sum_{k=0}^{\infty} (a_k^2 + b_k^2),$$

donde c_k, a_k y b_k son los coeficientes de Fourier de f.

14. Supongamos que f tiene una derivada segunda continua sobre $[-\pi, \pi]$ y que $f(-\pi) = f(\pi)$ y $f'(-\pi) = f'(\pi)$. Pruébese que

a) los coeficientes de Fourier, c_k, de f tienen la propiedad de que para algún $\alpha > 0$

$$|c_k| < \frac{\alpha}{k^2} \quad \text{para todo } k.$$

Sugerencia: intégrese por partes.

b) La serie de Fourier de f converge uniforme y absolutamente a f sobre $[-\pi, \pi]$.

c) Generalícese el resultado de la parte *a*) para cuando f tiene derivada m-ésima continua sobre $[-\pi, \pi]$.

15. Supongamos que f es continua sobre $[-\pi, \pi]$. Pruébese que:

a) para cada $\varepsilon > 0$ hay una función poligonal continua (su gráfica está formada de segmentos rectilíneos) p con la propiedad de que

$$|f(x)-p(x)| < \varepsilon \quad \text{para toda } x \text{ en } [-\pi, \pi].$$

b) Sea ε un número positivo cualquiera y $[a, b]$ un intervalo cerrado cualquiera en el interior de $[-\pi, \pi]$ $(-\pi < a < b < \pi)$. Sea p una función poligonal cualquiera continua sobre $[-\pi, \pi]$. Entonces existe una función g con una derivada segunda continua sobre $[-\pi, \pi]$ tal que $g(-\pi) = g(\pi)$, $g'(-\pi) = g'(\pi)$, y con la propiedad de que

$$|p(x)-g(x)| < \varepsilon \quad \text{para toda } x \text{ en } [a, b].$$

c) Para todo $\varepsilon > 0$ y todo intervalo $[a, b]$ en el interior de $[-\pi, \pi]$ hay un polinomio trigonométrico

$$T(x) = \sum_{k=-n}^{n} \alpha_k e^{ikx}$$

con la propiedad de que

$$|f(x) - T(x)| < \varepsilon \quad \text{para toda } x \text{ en } [a, b].$$

16. *Teorema de la aproximación de Weierstrass.* Si f es continua sobre $[a, b]$, entonces para todo $\varepsilon > 0$ hay un polinomio p con la propiedad de que

$$|f(x)-p(x)| < \varepsilon \quad \text{para toda } x \text{ en } [a, b].$$

Sugerencia: véase el problema 15.

9. APROXIMACIONES DE FOURIER

Hemos discutido el problema de la representación de una función por una serie trigonométrica. Las funciones exponenciales complejas y las funciones trigonométricas son nada más que dos elementos de la gran clase de las llamadas funciones "ortogonales".[1] Vamos a continuación a considerar el problema más general de aproximarnos a una función arbitraria por una combinación lineal de funciones "ortogonales". Estamos interesados en el problema de la aproximación en el espacio $\mathscr{C}$ de todas las

[1] Como el lector verá, no hay funciones ortogonales, sino "familias de funciones" ortogonales. Las funciones exponenciales complejas constituyen una familia de funciones ortogonales; las trigonométricas, otra. [N. del T.]

funciones complejas (con valores en el campo complejo) continuas sobre un intervalo $[a, b]$ de la recta real. Este espacio es completamente análogo a los espacios vectoriales R^n y C^n. La principal diferencia es que no es finitodimensional. Estamos ya familiarizados con las operaciones algebraicas sobre funciones y la única nueva característica en este espacio es la introducción de un producto interior o escalar de funciones análogo al producto escalar de vectores finito dimensionales. El espacio $\mathscr{C}$ es un espacio vectorial con un producto interior. Resolveremos el problema de la aproximación en un espacio vectorial V con un producto interior y, finalmente, aplicaremos esto a la aproximación de funciones por funciones ortogonales.

9.1 Definición. *Un **espacio vectorial** (complejo) es un conjunto de elementos $f, g, h, \ldots$, llamados vectores, con las siguientes propiedades (α, β son números complejos):*

A_1. *A cada par de vectores f y g de V corresponde un tercer vector h de V, llamado **suma** de f y g y escrito $h = f+g$.*

A_2. $f+g = g+f$.

A_3. $(f+g)+h = f+(g+h)$.

A_4. *Hay un vector (único) 0 en V, llamado **vector cero,** con la propiedad de que $f+0 = f$ para todo f de V.*

A_5. *Para cada f en V hay un vector (único), denotado por $-f$ y llamado **inverso** de f con la propiedad de que $f+(-f) = 0$.*

S_1. *Para cada número complejo α y cada vector f en V hay un vector h en V, escrito $h = \alpha f$.*

S_2. $1f = f$.

S_3. $\alpha(\beta f) = (\alpha\beta)f$.

S_4. $(\alpha+\beta)f = \alpha f+\beta f$.

S_5. $\alpha(f+g) = \alpha f+\alpha g$.

9.2 Definición. *Un **producto interior** o **escalar** sobre un espacio vectorial V, escrito (f, g), es una función compleja sobre $V \times V$ con las siguientes propiedades:*

9.3 $$(f, g) = \overline{(g, f)}$$

9.4 $$(\alpha f, g) = \alpha(f, g)$$

9.5 $$(f+g, h) = (f, h)+(g, h)$$

9.6 $$(f, f) \geqslant 0; \quad (f, f) = 0 \quad \text{si y sólo si } f = 0.$$

Nótese que por 9.3 y 9.4

9.7 $$(f, \alpha g) = \bar{\alpha}(f, g).$$

9.8 Definición. *Un conjunto finito de vectores $f_1, f_2, \ldots, f_n$ se dice que es* ***linealmente independiente*** *si $\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_n f_n = 0$ implica $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. En cualquier otro caso el conjunto se dice que es* ***linealmente dependiente*** *(es decir, existen números α_i, no todos cero, que satisfacen $\sum_{k=1}^{n} \alpha_i f_i = 0$). Los vectores f, g se dice que son* ***ortogonales*** *si $(f, g) = 0$. El número real $\|f\| = (f, f)^{1/2}$ se llama* ***norma*** *o* ***longitud*** *de f.*

Supongamos que f y g son ortogonales. Entonces tenemos

$$\|f+g\|^2 = (f+g, f+g) = (f, f)+(g, f)+(f, g)+(g, g)$$

y de aquí, como $(g, f) = \overline{(f, g)} = 0$,

9.9
$$\|f+g\|^2 = \|f\|^2+\|g\|^2.$$

Lo que se corresponde con el teorema de Pitágoras.

Sean f y g un par de vectores cualesquiera con $f \neq 0$. Entonces $h = g - \frac{(f, g)}{(f, f)} f$ es ortogonal a f: $(f, h) = (f, g) - \frac{(f, g)}{(f, f)}(f, f) = 0$.

Por tanto,

$$(g, h) = \left(h + \frac{(f, g)}{(f, f)} f, h\right) = (h, h) = \|h\|^2.$$

Continuando tenemos

$$0 \leqslant \|h\|^2 = (g, h) = \left(g, g - \frac{(f, g)}{(f, f)} f\right) = (g, g) - \frac{(f, g)(g, f)}{(f, f)}.$$

De donde

$$0 \leqslant (f, f)\, \|h\|^2 = (f, f)(g, g) - |(f, g)|^2.$$

Esta desigualdad se verifica también si $f = 0$, y es válida, por tanto, para todo f. Como claramente $(f, f)\, \|h\|^2 = 0$ si y sólo si f y g son linealmente dependientes, hemos probado la *desigualdad de Schwarz*

9.10
$$|(f, g)| \leqslant \|f\|\, \|g\|$$

donde se tiene la igualdad si y sólo si f y g son linealmente dependientes.

Nuestro interés principal recae en un problema de aproximación, y este problema adquiere su significación más alta en espacios que no son finitodimensionales. Tiene, sin embargo, un sencillo significado geométrico en espacios finitodimensionales. Sean $\mathbf{u}_1$ y $\mathbf{u}_2$ un par de vectores ortogonales unitarios en R^3 (figura 3), es decir, $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$, $\mathbf{u}_1 \cdot \mathbf{u}_1 = 1$, $\mathbf{u}_2 \cdot \mathbf{u}_2 = 1$. Sea $\mathbf{x}$ otro vector cualquiera en R^3. Queremos aproximarnos a $\mathbf{x}$ por una combinación lineal $\mathbf{y} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2$ de $\mathbf{u}_1$ y $\mathbf{u}_2$. La aproximación ha de

encontrarse, pues, en el plano $\mathscr{P}$ que pasa por el origen y está determinado por los vectores $\mathbf{u}_1$ y $\mathbf{u}_2$. Tomamos como error de la aproximación $E = |\mathbf{x}-\mathbf{y}| = ((\mathbf{x}-\mathbf{y}) \cdot (\mathbf{x}-\mathbf{y}))^{1/2}$. Esta es la "raíz del error cuadrático",[1] y la "mejor" aproximación $\mathbf{y}$ es la que minimiza este error. Esta distancia $|\mathbf{x}-\mathbf{y}|$ se minimiza tomando como aproximación $\mathbf{y}$ la proyección ortogonal de $\mathbf{x}$ sobre el plano $\mathscr{P}$. Por tanto, la "mejor" aproximación $\mathbf{y}$ se obtiene tomando $\alpha_i = (\mathbf{x} \cdot \mathbf{u}_i)$.

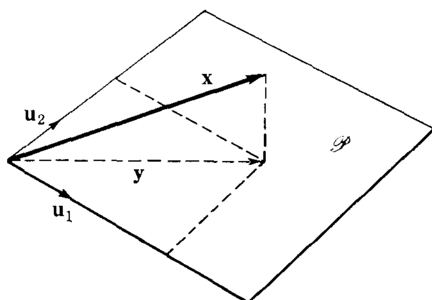


FIGURA 3

La extensión de este resultado a cualquier espacio vectorial V es directa. Una sucesión de vectores $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots$ en V se dice que es *ortonormal* si tienen longitud unidad y son ortogonales a pares; es decir, $(\varphi_j, \varphi_k) = 0$ si $j \neq k$ y $(\varphi_j, \varphi_j) = 1$ para cualesquier $j, k = 1, 2, \ldots$ Sea f un vector arbitrario en V. Queremos aproximar f por una combinación lineal $g_n = \sum_{k=1}^{n} \alpha_k \varphi_k$ de los primeros n vectores de una sucesión ortonormal $\varphi_1, \varphi_2, \ldots$ El *error* de una aproximación g_n es $E_n = (f-g_n, f-g_n)^{1/2} = \|f-g_n\|$. La *mejor aproximación* g_n es aquella que minimiza el error E_n.

9.11 Teorema. *Si $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots$ es una sucesión ortonormal, entonces la mejor aproximación $\sum_{k=1}^{n} \alpha_k \varphi_k$ de f está dada tomando*

9.12 $$\alpha_k = (f, \varphi_k).$$

El error cuadrático mínimo es

9.13 $$E_n^2(f) = \|f\|^2 - \sum_{k=1}^{n} |\alpha_k|^2 .$$

[1] Si en un espacio vectorial hemos definido una norma $\| \;\|$, ella nos induce una distancia: $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|$. Estamos, pues, tomando como error la distancia entre el vector $\mathbf{x}$ y la aproximación $\mathbf{y}$. [N. del T.]

PRUEBA. Calculando el error cuadrático, tenemos

$$E_n^2 = \left(f - \sum_{k=1}^{n} \alpha_k \varphi_k, f - \sum_{k=1}^{n} \alpha_k \varphi_k\right)$$

$$= (f, f) - \sum_{k=1}^{n} \bar{\alpha}_k (f, \varphi_k) - \sum_{k=1}^{n} \alpha_k (\varphi_k, f) + \sum_{k=1}^{n} \sum_{j=1}^{n} \alpha_k \bar{\alpha}_j (\varphi_k, \varphi_j).$$

Haciendo

$$\beta_k = (f, \varphi_k),$$

y recordando que las φ_k son ortonormales, obtenemos

$$E_n^2 = (f, f) - \sum_{k=1}^{n} \bar{\alpha}_k \beta_k - \sum_{k=1}^{n} \alpha_k \bar{\beta}_k + \sum_{k=1}^{n} \alpha_k \bar{\alpha}_k$$

$$= \|f\|^2 + \sum |\beta_k - \alpha_k|^2 - \sum \beta_k \bar{\beta}_k .$$

Es entonces evidente que la elección $\alpha_k = \beta_k = (f, \varphi_k)$ minimiza el error, y el error cuadrático mínimo es

$$E_n^2(f) = \|f\|^2 - \sum_{k=1}^{n} |\beta_k|^2 .$$

Los coeficientes $\alpha_k = (f, \varphi_k)$ se llaman *coeficientes de Fourier* de f con respecto a la sucesión ortonormal $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots$ La aproximación $\sum_{k=1}^{n} \alpha_k \varphi_k$ se llama *n-ésima aproximación de Fourier*. Como $E_n^2(f) \geqslant 0$ y es no creciente $\lim_{n\to\infty} E_n^2(f)$ existe y, por tanto, también $\sum_{k=1}^{\infty} |\alpha_k|^2$. De donde tenemos

$$0 \leqslant \lim_{n\to\infty} E_n^2(f) = \|f\|^2 - \sum_{k=1}^{\infty} |\alpha_k|^2$$

y

9.14 $$\sum_{k=1}^{\infty} |\alpha_k|^2 \leqslant \|f\|^2 .$$

Desigualdad a la que se llama *desigualdad de Bessel.*

Aumentando el número de términos en la aproximación nos gustaría ser capaces de reducir el error en la aproximación si, desde luego, el error no es ya cero. Una propiedad de los sistemas ortonormales que nos asegura esto es la complitud: una sucesión ortonormal $\{\varphi_k\}$ se dice que es *completa* si no existe ningún vector distinto de cero en V ortogonal a todas y cada una de las φ_k; es decir $(h, \varphi_k) = 0$ para toda $k = 1, 2, \ldots$ implica $h = 0$.

Por 9.13 vemos que el error puede siempre reducirse a menos que $\alpha_k = (f, \varphi_k) = 0$ para todo $k > n$; es decir, a menos que $h = f - \sum_{k=1}^{n} \alpha_k \varphi_k$

sea ortogonal a toda φ_k, $k = 1, 2, \dots$ De donde para una sucesión completa o la aproximación n-ésima es exacta ($h = 0$) o el error puede reducirse añadiendo más términos.

Una propiedad aún más conveniente es la de $\lim_{n\to\infty} E_n(f) = 0$ para todo f en V. Una sucesión ortonormal con esta propiedad se dice que es *cerrada.* Con un sistema ortonormal cerrado, el error en la aproximación puede hacerse tan pequeño como se quiera tomando un número suficiente de términos en la aproximación de Fourier. Es claro que si una sucesión ortonormal es cerrada entonces es también completa, pues, ciertamente no puede ser cerrada si existe un vector distinto de cero h en V todos los coeficientes de Fourier del cual son cero.

Veamos ahora cómo se aplica esto a la aproximación de funciones. Sea $\mathscr{C}$ el espacio de todas las funciones complejas continuas sobre un intervalo $[a, b]$ de la recta real. Para funciones f, g en $\mathscr{C}$ y cualquier número complejo α, $f+g$ se definen en la forma usual. Definimos también

9.15
$$(f, g) = \int_a^b f\bar{g}\,.$$

Puede entonces verificarse (problema 1) que $\mathscr{C}$ es un espacio vectorial y (f, g) un producto interior.[1] Los siguientes son ejemplos de sucesiones ortonormales en $\mathscr{C}$:

9.16 Las funciones exponenciales complejas normalizadas $\dfrac{1}{\sqrt{2\pi}}$, $\dfrac{1}{\sqrt{2\pi}}e^{ix}$, $\dfrac{1}{\sqrt{2\pi}}e^{-ix}$, $\dfrac{1}{\sqrt{2\pi}}e^{2ix}$, $\dfrac{1}{\sqrt{2\pi}}e^{-2ix}$, $\cdots$, forman una sucesión ortonormal sobre el intervalo $[-\pi, \pi]$.

9.17 Las funciones trigonométricas normalizadas $\dfrac{1}{\sqrt{\pi}}$, $\dfrac{1}{\sqrt{\pi}}\cos x$, $\dfrac{1}{\sqrt{\pi}}\operatorname{sen} x$, $\dfrac{1}{\sqrt{\pi}}\cos 2x$, $\dfrac{1}{\sqrt{\pi}}\operatorname{sen} 2x$, $\cdots$, forman una sucesión ortonormal sobre el intervalo $[-\pi, \pi]$.

9.18 Las funciones trigonométricas normalizadas $\sqrt{\dfrac{2}{\pi}}\operatorname{sen} x$, $\sqrt{\dfrac{2}{\pi}}\operatorname{sen} 2x$, $\cdots$ forman una sucesión ortonormal sobre el intervalo $[0, \pi]$.

[1] Entonces $\|f\|^2 = \int_a^b |f|^2$. En la sección 2, pág. 698, consideramos una norma diferente y más "fuerte", y no debemos confundir una con otra. La norma de la sección 2 se dice que es más "fuerte" porque la convergencia según esa norma (que es convergencia uniforme sobre $[a, b]$) implica la convergencia según esta última que hemos definido.

9.19 Las funciones trigonométricas normalizadas $\frac{1}{\sqrt{\pi}}$, $\sqrt{\frac{2}{\pi}}\cos x$, $\sqrt{\frac{2}{\pi}}\cos 2x$, forman una sucesión ortonormal sobre el intervalo $[0, \pi]$.

9.20 Los polinomios de Legendre normalizados

$$\varphi_n(x) = \sqrt{\frac{2n+1}{2}}\, P_n(x), \quad n = 0,1.2,\cdots$$

forman una sucesión ortonormal sobre el intervalo $[-1, 1]$.

En la sección 8 los coeficientes de Fourier son los de las funciones trigonométricas y exponenciales y aquí, en esta sección, los coeficientes de Fourier son coeficientes de las funciones normalizadas. Ambos conducen a las mismas aproximaciones de f y si s_n es la n-ésima aproximación son ellos los que minimizan el error cuadrático

$$E_n^2(f) = \int_a^b |f - s_n|^2 .$$

Para algunos propósitos, el error cuadrático medio $e_n^2 = (b-a)^{-1} E_n^2$ es el que se usa, y la *raíz del error cuadrático medio* es $e_n = (b-a)^{-1/2} E_n$. Los coeficientes de Fourier minimizan e_n al igual que E_n y se dice que dan la mejor aproximación cuadrática media.

9.21 Ejemplo. Calcúlese la raíz del error cuadrático medio en el ejemplo 8.18 usando los primeros tres términos distintos de cero de la serie de Fourier como aproximación.

SOLUCIÓN. El factor normalizante para la sucesión ortogonal sen $k\pi\frac{x}{l}$, $k = 1, 2, \cdots$, sobre el intervalo $[0, l]$ es $\sqrt{\frac{2}{l}}$. Por tanto, los coeficientes b_k del ejemplo 8.18 y los α_k en 9.12 están relacionadas por $b_k = \sqrt{\frac{2}{l}}\,\alpha_k$. De donde

$$E_5^2(f) = \int_0^l f^2 - \frac{64}{\pi^4}\,\frac{l}{2}\left(1 + \frac{1}{3^4} + \frac{1}{5^4}\right).$$

Ahora

$$\int_0^l f^2 = 2\int_0^{\frac{1}{2}l} \left(\frac{2}{l}x\right)^2 dx = \frac{1}{3}l$$

y

$$E_5^2(f) = l[\tfrac{1}{3} - 0.33308] = 0.00025\,l.$$

La raíz del error cuadrático medio es entonces

$$e_5 = 0.016.$$

9.22 Ejemplo. Determínese la mejor aproximación cuadrática media de $|x|$ por un polinomio de grado 4 sobre el intervalo $[-1, 1]$. Calcúlese la raíz del error cuadrático medio.

SOLUCIÓN. Los polinomios de Legendre normalizados son

$$\varphi_n(x) = \sqrt{\frac{2n+1}{2}}\, P_n(x).$$

Como todo polinomio de grado n puede expresarse como una combinación lineal de los primeros n polinomios de Legendre, se sigue que la mejor aproximación cuadrática media de $|x|$ de grado n sobre $[-1, 1]$ es

$$s_n(x) = \sum_{k=0}^{n} \alpha_k \varphi_k(x)$$

donde

$$\alpha_k = \int_{-1}^{1} |x|\, \varphi_k(x)\, dx.$$

Como $P_k(x)$ es impar si k es impar y par si k es par,

$$\alpha_{2k+1} = 0$$

y

$$\alpha_{2k} = 2\int_0^1 x\varphi_{2k}(x)\, dx.$$

Ahora

$$\varphi_0(x) = \frac{1}{\sqrt{2}}, \qquad \varphi_2(x) = \frac{1}{2}\sqrt{\frac{5}{2}}(3x^2-1)$$

$$\varphi_4(x) = \frac{3}{8\sqrt{2}}(35x^4-30x^2+3).$$

De donde

$$\alpha_0 = \frac{1}{\sqrt{2}}, \qquad \alpha_2 = \frac{1}{4}\sqrt{\frac{5}{2}}, \qquad \alpha_4 = -\frac{1}{8\sqrt{2}},$$

y la mejor aproximación cuadrática media de grado 4 es

$$s_4(x) = \frac{1}{2} + \frac{5}{16}(3x^2 - 1) - \frac{3}{128}(35x^4 - 30x^2 + 3)$$

$$= \frac{1}{128}(-105x^4 + 210x^2 + 15).$$

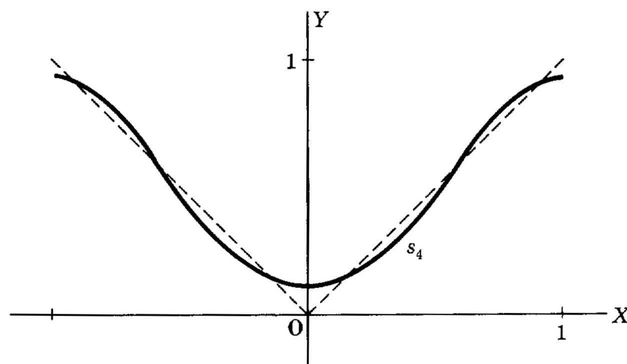


FIGURA 4

Esta aproximación polinomial está dibujada en la figura 4.

Obtenemos entonces

$$E_4{}^2 = \int_{-1}^{1} |x|^2\, dx - \left(\frac{1}{2} + \frac{5}{32} + \frac{1}{128}\right)$$

$$= \frac{2}{3} - \frac{85}{128} = \frac{1}{384}$$

y

$$e_4 = \frac{1}{\sqrt{768}} = \frac{1}{16\sqrt{3}}.$$

Todas las sucesiones de funciones ortonormales que hemos mencionado hasta el momento se sabe que son cerradas (problemas 7 y 8) y son, por tanto, completas. El teorema 8.8 fue una prueba de la complitud de la sucesión exponencial 9.16 y, como consecuencia, una prueba de la complitud de las sucesiones trigonométricas 9.17, 9.18 y 9.19. Estas sucesiones ortonormales son, todas, soluciones de una clase especial de ecuaciones diferenciales lineales de segundo orden, y la teoría de tales sucesiones ortonormales se llama teoría de Sturm-Liouville (referencias [58] y [56]).

Problemas

1. Pruébese que: *a*) el espacio $\mathscr{C}$ de todas las funciones continuas complejas sobre un intervalo $[a, b]$ es un espacio vectorial. *b*) $(f, g) = \int_a^b f\bar{g}$ es un producto interior de $\mathscr{C}$.

2. Problema 12, sección 8.

3. Problema 13, sección 8.

4. Obténgase la mejor aproximación cuadrática media de f sobre $[-1, 1]$ por un polinomio de grado 3 y calcúlese la raíz del error cuadrático medio de la aproximación.

a) $f(x) = e^x$
b) $f(x) = |x|^{1/2}$
c) $f(x) = \dfrac{1}{1+x^2}$.

5. Calcúlese la raíz del error cuadrático medio de la aproximación de Fourier de

a) $f(x) = x$ por $\sum_{k=1}^{n} b_k \operatorname{sen} kx$ sobre $[0, \pi]$, $n = 1, 2, 3, 4$

b) $f(x) = 1$ por $\sum_{k=1}^{n} b_k \operatorname{sen} k\pi x$ sobre $[0, 1]$, $n = 1, 3, 5, 7$

c) $f(x) = x^2$ por $\frac{1}{2}a_0 + \sum_{k=1}^{n} a_k \cos kx$ sobre $[0, \pi]$, $n = 1, 2, 3, 4$

d) $f(x) = 0$ para $x \leqslant 0$ y $f(x) = x$ para $x > 0$ sobre $[-\pi, \pi]$ por $\frac{1}{2}a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \operatorname{sen} kx)$ para $n = 0, 1, 2, 3, 4$.

6. Determínese la mejor aproximación cuadrática media de seno por un polinomio de grado 5 sobre el intervalo

a) $[-1, 1]$
b) $[-\pi/2, \pi/2]$
c) $[0, \pi]$.

7. Pruébese que las sucesiones trigonométricas 9.17, 9.18 y 9.19 son sucesiones ortonormales cerradas. *Sugerencia:* véase el problema 15*c*, sección 8.

8. Pruébese que la sucesión de polinomios de Legendre normalizados de 9.20 es una sucesión ortonormal cerrada. *Sugerencia:* véase el problema 16, sección 8 y 7.11.

9. Sea $\{\varphi_k\}$ una sucesión ortonormal en un espacio vectorial V. Sean $\{\alpha_k\}$ los coeficientes de Fourier de f y $\{\beta_k\}$ los coeficientes de Fourier de g. Pruébese que $\{\varphi_k\}$ *es cerrada si y sólo si*

$$(f, g) = \sum_{k=0}^{\infty} \alpha_k \bar{\beta}_k$$

para toda f, g en V.

Sugerencia:

$$4(f, g) = \|f+g\|^2 - \|f-g\|^2 + i\|f+ig\|^2 - i\|f-ig\|^2.$$

10. El espacio l^2 es el conjunto de todas las sucesiones complejas $a = \{a_n\}$ con la propiedad de que $\sum_{k=0}^{\infty} |a_k|^2$ converja. Pruébese que con la definición habitual de adición de sucesiones y de multiplicación de una sucesión por un complejo, l^2 es un espacio vectorial. Para $a = \{a_n\}$, $b = \{b_n\}$ definamos $(a, b) = \sum_{k=0}^{\infty} a_k \bar{b}_k$. Pruébese que (a, b) es un producto interior sobre l^2.

11. Sea $\{\varphi_k\}$ una sucesión ortonormal en un espacio vectorial, V, con producto interior. Para cada f en V definamos

$$Q(f) = \{\alpha_k\}$$

donde $\{\alpha_k\}$ es la sucesión de coeficientes de Fourier $\alpha_k = (f, \varphi_k)$. Q es entonces una función de V en l^2. Esta función Q puede pensarse como reemplazando a la asociación formal $f \sim \sum_{k=0}^{\infty} \alpha_k \varphi_k$ de f con su serie de Fourier. La sucesión ortonormal $\{\varphi_k\}$ es análoga a una base de V y las α_k son como coordenadas cartesianas. Derívense algunas propiedades de la función Q. En este contexto, ¿cuál es el significado de la hipótesis de que 1) $\{\varphi_k\}$ es completa?, 2) $\{\varphi_k\}$ ¿es cerrada?

Nota. No es, en general, cierto que toda sucesión en l^2 es una sucesión de coeficientes de Fourier de alguna f en V. El rango de Q puede no ser l^2. Un teorema ahora famoso —llamado teorema de Riesz-Fischer— afirma que para algunos espacios (espacios L^2) toda sucesión en l^2 es una sucesión de coeficientes de Fourier de alguna función en el espacio. Este importante resultado fue descubierto casi simultáneamente por Friedrich Riesz y Ernst Fischer en 1907.

Respuestas a problemas escogidos

Páginas 19-20

1. *a*) (5, 0, 11); *c*) (17, −1, 37); *e*) $(-t, 5-10t, 6-4t)$;
g) $t = -3$, (−9, 20, −6); $t = -2$, (−6, 15, −2);
$t = -1$, (−3, 10, 2); $t = 0$, (0, 5, 6).
6. *a*) $\frac{1}{8}(1, -13)$; *c*) (22, −3, −12, −19).
7. *a*) No hay soluciones; *c*) no hay soluciones; *e*) todo r real.
8. *a*) $r = s = 0$; *c*) $r = s = 0$; *e*) $r = 6$, $s = -7$.

Páginas 23-24

1. *a*) (8, −8); *c*) (−4, 4); *e*) $\frac{1}{2}(\sqrt{3}+\sqrt{2}, 1+\sqrt{2})$.
2. *a*) $r = \frac{5}{11}$, $s = \frac{10}{11}$; *c*) $r = -\frac{13}{5}$, $s = -\frac{7}{5}$; *e*) $r = s = 2$.
3. $\mathbf{a}+\mathbf{b}+\mathbf{c} = \mathbf{0}$.

Página 24

1. *a*) Igual dirección; *c*) no paralelas; *e*) dirección opuesta.

Página 28

1. *a*) $\sqrt{34}$; *c*) $\sqrt{30}$; *e*) $3\sqrt{34}$; *g*) $\frac{1}{2}\sqrt{14}$; *i*) 1; *k*) 1.
4. *a*) Ortogonal; *c*) ortogonal.
7. *a*) $r(-2, 1), r \in R$; *c*) $r(-a_2, a_1), r \in R$.
9. *a*) $\dfrac{\sqrt{2}}{2}(1, 1)$; *c*) $\dfrac{1}{\sqrt{62}}(2, 3, -7)$.

Página 31

1. *a*) −9; *c*) 3; *e*) 34; *g*) 20.
2. *a*) Ortogonal; *c*) no ortogonal.
3. *a*) $r(-2, 1), r \in R$; *c*) $r(0, 0, 1), r \in R$; *e*) $r(-a_1, a_2), r \in R$.

Páginas 35-36

1. *a*) $3(1, 0)+8(0, 1)$; *c*) $\frac{3}{2}(1, 1)+\frac{13}{2}(-1, 1)$; *e*) $\frac{3}{2}(1, 1, 0)+\frac{1}{2}(-1, 1, 6)$
3. *a*) $\text{Comp}_{\mathbf{b}}\ \mathbf{a} = 3$, $\mathbf{Proy}_{\mathbf{b}}\ \mathbf{a} = (3, 0)$;
c) $\text{Comp}_{\mathbf{b}}\ \mathbf{a} = -3$, $\mathbf{Proy}_{\mathbf{b}}\ \mathbf{a} = (0, 0, -3)$;
e) $\text{Comp}_{\mathbf{b}}\ \mathbf{a} = \frac{2}{3}\sqrt{3}$, $\mathbf{Proy}_{\mathbf{b}}\ \mathbf{a} = \frac{2}{3}(1, 1, 1)$;
g) $\text{Comp}_{\mathbf{b}}\ \mathbf{a} = |a_2|$, $\mathbf{Proy}_{\mathbf{b}}\ \mathbf{a} = \mathbf{b}$.

Páginas 38-39

1. *a*) $(-2, 3, 9, 0)$; *b*) 0; *c*) 18; *d*) $\sqrt{50}$;
e) $\frac{1}{\sqrt{63}}(-2, 3, -1, 7)$; *f*) $(1, \frac{11}{2}, \frac{11}{2}, -\frac{1}{2})$.
2. *a*) Dirección opuesta; *b*) no paralela; *c*) igual dirección;
d) no paralela.
3. *a*) $\sqrt{46}$; *b*) $\sqrt{15}$; *c*) $\sqrt{47}$; *d*) $\sqrt{75}$.
4. *a*) Ortogonal; *b*) ortogonal; *c*) no ortogonal; *d*) ortogonal.
5. *a*) $\text{Comp}_{\mathbf{b}}\ \mathbf{a} = \frac{2}{3}\sqrt{6}$, $\mathbf{Proy}_{\mathbf{b}}\ \mathbf{a} = \frac{2}{3}(-1, 2, 1)$;
b) $\text{Comp}_{\mathbf{b}}\ \mathbf{a} = \frac{5}{6}\sqrt{2}$, $\mathbf{Proy}_{\mathbf{b}}\ \mathbf{a} = \frac{5}{18}(-1, 3, -2, 2)$;
c) $\text{Comp}_{\mathbf{b}}\ \mathbf{a} = 0$, $\mathbf{Proy}_{\mathbf{b}}\ \mathbf{a} = \mathbf{0}$;
d) $\text{Comp}_{\mathbf{b}}\ \mathbf{a} = \frac{1}{2}\sqrt{2}$, $\mathbf{Proy}_{\mathbf{b}}\ \mathbf{a} = \frac{1}{2}(-1, 0, 1)$.

Páginas 47-48

1. *a*) $\sqrt{35}$; *c*) $\sqrt{91}$; *e*) $\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}$; *g*) $|t|\ |\mathbf{a}|$.
2. *a*) $\{t(1, 1, 1)\}$; *c*) $\{(7, 12, -11)+t(0, 0, 1)\}$;
e) $\{(-3, 2, -1)+t(1, 5, -4)\}$.
3. *a*) $\{t(1, 1, 1)\}$ *c*) $\{(8, -3, 2)+t(-3, 3, -2)\}$;
e) $\{(-3, 2, -1)+t(1, 5, -4)\}$.
4. *a*) $\{(2, 3, 1)+u(1, 2, 5)+v(5, -1, 3)\}$; *c*) $\{u(1, 1, 1)+v(2, 0, 0)\}$;
e) $\{(1, 1, 1)+u(2, -3, -1)+v(3, 2, -2)\}$.
5. *a*) Sí; *c*) no.
7. *a*) $\{(3, 4, 5)\}$; *c*) $\{\frac{73}{18}(1, 1, 1)\}$; *e*) $\{\frac{1}{13}(-58, -69, 63)\}$.

Páginas 53-54

1. *a*) No paralelas, $\varnothing$; *c*) no paralelas, $\varnothing$;
e) no paralelas $\{(\frac{3}{5}, -\frac{4}{5}, \frac{17}{5})\}$.
3. *a*) $\alpha = \beta = \gamma = \text{arccos}\,\frac{1}{\sqrt{3}} = 54°\,44'$; *c*) $\theta = \text{arccos}\,\frac{1}{3} = 70°\,32'$.
5. *a*) $\{t(\sqrt{2}, 1, \pm 1)\}$; *c*) $\{t(1, 1, 1)\}$.

Páginas 57-58-59

1. *a)* $(-42, 13, 59)$; *c)* $(-42, -73, 16)$; *e)* $(-252, 78, 354)$;
g) 903; *i)* 0; *k)* $(-134, -25, 155)$.
4. *a)* $r(0, 0, 1), r \in R$; *c)* $r(-41, -18, 7), r \in R$;
e) $r(2, 0, 1)+s(-3, 1, 0), r, s \in R$.
9. *a)* 26; *c)* $\sqrt{76{,}889}$; *e)* $3\sqrt{14}$.
10. *a)* 4; *c)* $\sqrt{6361}$; *e)* $\frac{1}{2}\sqrt{3}$.

Página 62

3. *a)* 96; *c)* 20; *e)* 46. **4.** *a)* $\frac{4}{3}$; *c)* $\frac{45}{2}$; *e)* $\frac{11}{3}$.

Páginas 65-66-67

1. *a)* Dependientes; *c)* independientes. **3.** *a)* $-2, 0, 2$.

Página 72

1. *a)* $(24, 9, -29)$; *c)* $(2, 1, -9)$; *e)* $-3(14, 3, -30)$.
2. *a)* $x+y+z = 0$; *c)* $2x+9y-3z = -43$;
e) $(\mathbf{P}_2-\mathbf{P}_1)\cdot\mathbf{P} = \frac{1}{2}(\mathbf{P}_2-\mathbf{P}_1)\cdot(\mathbf{P}_2+\mathbf{P}_1)$.
3. *a)* $24x+9y-29z = 164$; *c)* $2x+y-9z = 29$;
e) $14x+3y-30z = -65$.
5. *Sugerencia:* pruébese que $|\mathbf{P}-\mathbf{P}_1| = |\mathbf{P}-\mathbf{P}_2|$ si y sólo si $\mathbf{n}\cdot(\mathbf{P}-\mathbf{P}_0) = 0$ donde $\mathbf{P}_0 = \frac{1}{2}(\mathbf{P}_1+\mathbf{P}_2)$ y $\mathbf{n} = \mathbf{P}_2-\mathbf{P}_1$.

Página 75

1. *a)* $\{t(10, -7, 7)\}$; *c)* $\{(\frac{38}{57}, -1, 0)+t(8, 99, 57)\}$;
e) $\{\frac{1}{2}(1, 1, 0)\}$; *g)* $\{\frac{1}{20}(-29, 44, 39)\}$.
2. *a)* $\operatorname{arccos}\dfrac{-6}{\sqrt{234}} = 113^\circ 5'$; *c)* $\operatorname{arccos}\dfrac{82}{\sqrt{19838}} = 54^\circ 24'$.

3. $\operatorname{arccos}\dfrac{-1}{\sqrt{105}} = 95^\circ 36'$.

Páginas 77-78

1. *a)* $\{\frac{1}{4}(-1, -5, 7)\}$, no paralela; *c)* $\varnothing$, paralela;
e) $\{\frac{1}{3}(1, -2, 1)\}$, no paralela.

2. *a*) $\{(1, 2, 3)+t(1, 1, -2)\}$; *c*) $\{(0, 2, -2)+t(7, -1, -7)\}$.
3. *a*) $\frac{4}{3}\sqrt{6}$; *c*) $\frac{8}{33}\sqrt{11}$. **5.** $\frac{1}{38}(77, 88, 103)$.
7. $(1, 10, 4)$, $\frac{50}{39}\sqrt{39}$. **9.** $\frac{1}{7}(11, 2, 0)$.

Páginas 84-85

1. *a*) $\mathbf{d} = 3\mathbf{a}+\mathbf{b}-3\mathbf{c}$; *c*) $\mathbf{d} = 2\mathbf{a}-\frac{5}{7}\mathbf{b}+\frac{3}{7}\mathbf{c}$.
3. *a*) $(1, 2, 1)$, $\frac{7}{3}(1, -1, 1)$, $\frac{1}{2}(1, 0, -1)$; *c*) $\mathbf{b}_1 = \mathbf{b}$, $\mathbf{c}_1 = \mathbf{c}$.
4. *a*) $\mathbf{d} = -\frac{3}{7}\mathbf{a}-\frac{23}{7}\mathbf{b}+5\mathbf{c}$. **5.** *a*) $(\frac{3}{2}, 0, \frac{1}{2})$.

Páginas 88-89

1. *a*) $(0, 1, 0)$; *c*) $(0, -3, 10)$; *e*) $(0, -1, -3)$.
2. *a*) $(1, 0, 0)$; *c*) $\frac{1}{2}(0, \sqrt{2}, \sqrt{2})$; *e*) $\frac{1}{4}(-13, 13\sqrt{3}, 26\sqrt{3})$.
3. *a*) $(r, \theta, z) = (1, 0, 0)$, $(\rho, \theta, \varphi) = (1, 0, \pi/2)$;
c) $(r, \theta, z) = (0, \theta, 1)$, $(\rho, \theta, \varphi) = (1, \theta, 0)$, θ arbitrario;
e) $(r, \theta, z) = (0, \theta, 0)$, $(\rho, \theta, \varphi) = (0, \theta, \varphi)$, θ y φ arbitrario;
g) $(r, \theta, z) = (\sqrt{1649}, -38°, 18)$, $(\rho, \theta, \varphi) = (\sqrt{1973}, -38°, 66°6')$.

Página 94

1. *a*) $\mathscr{L} = \{(1, 1, 1, 1)+t(0, 1, 2, 3)\}$, $k = 1$;
c) $\mathscr{P} = \{(1, 0, 2, 0)+t_1(1, 0, 1, 1)+t_2(0, 1, 1, 1)\}$, $k = 2$.
2. *a*) $\sqrt{39}$; *c*) $\sqrt{26}$. **3.** $\mathscr{P}_1 \cap \mathscr{P}_2 = \varnothing$.

Páginas 95-96

1. *a*) $\{(1, 1, 1)+t(3, 0, 4)\}$; *b*) $(\mathbf{P}_1-\mathbf{P}_2)\times(\mathbf{P}_1-\mathbf{P}_3) = (-12, 1, 9) \neq \mathbf{0}$;
c) $12x-y-9z = 2$; *d*) $\frac{1}{2}\sqrt{226}$; *e*) $\{(-5, 2, -1)+t(12, -1, -9)\}$;
f) $\frac{55}{226}\sqrt{226}$; *g*) $\{\frac{1}{12}(17, -9, 21)\}$; *h*) $\{(-5, 2, -1)+t(3, 0, 4)\}$;
i) $119°11'$; *j*) $\frac{1}{\sqrt{226}}(12, -1, -9)$.
2. *a*) $\sqrt{35}$; *b*) $\frac{1}{2}\sqrt{3}$; *c*) 2; *d*) $\frac{1}{3}$. **4.** 144. **5.** $48°52'$.

Páginas 100-101

1. *a*) $\mathbf{f}(t) = (-1, 2)+t(4, 3)$; *c*) $\mathbf{f}(t) = (2t+1, -6t+4, -6t+7)$.

Páginas 107-108

3. *a*) $(\sqrt{2}, 4, \text{sen } 2)$; *c*) no existe límite. **9.** $(1, 2t, 3t^2)$.

Página 110

1. *a*) Ninguno; *c*) 1, 2, 3, 4.

Página 114

1. *a*) Espiral; *c*) segmento rectilíneo.

5. $\mathbf{f}(t) = \left(\dfrac{t}{\sqrt{1+t^2}}, \dfrac{t^2}{\sqrt{1+t^2}}\right)$.

6. *a*) $\{(t, -\frac{1}{2}gt^2+2t) \mid t \in \mathrm{R}\}$; *c*) $\{(e^t, e^t+e^{2t}) \mid t \in \mathrm{R}\}$.

Páginas 121-122

1. *a*) $\mathbf{f}' = (\frac{1}{3}I^{-2/3}, 6I, \cos)$; *c*) $\mathbf{f}'(t) = \left(\dfrac{2t}{t^2+1}, \dfrac{t}{\sqrt{t^2+1}}, \dfrac{2-2t^2}{(t^2+1)^2}\right)$;

e) $\mathbf{f}'(t) = (1, 1, 0)$, t no un entero.

3. *a*) en $(0, 3)$: $(-4, 0)$, $\{(0, 3)+t(-4, 0)\}$; en $2(\sqrt{2}, \frac{3}{2}\sqrt{2})$: $(-2\sqrt{2}, \frac{3}{2}\sqrt{2})$, $\{(2\sqrt{2}, \frac{3}{2}\sqrt{2})+t(-2\sqrt{2}, \frac{3}{2}\sqrt{2})\}$; en $(4, 0)$: $(0, 3)$, $\{(4, 0)+t(0, 3)\}$;

c) en $(0, 0, 0)$: $\left(1, 0, \dfrac{1}{2\pi}\right)$; $\left\{t\left(1, 0, \dfrac{1}{2\pi}\right)\right\}$; en $\left(0, \dfrac{\pi}{2}, \dfrac{1}{4}\right)$: $\left(-\dfrac{\pi}{2}, 1, \dfrac{1}{2\pi}\right)$;

$\left\{\left(0, \dfrac{\pi}{2}, \dfrac{1}{4}\right) + t\left(-\dfrac{\pi}{2}, 1, \dfrac{1}{2\pi}\right)\right\}$.

4. *a*) Tangente horizontal en $(-1, -1)$; tangente vertical en $(0, 0)$;

c) tangente horizontal en $\left(\dfrac{\sqrt{3}}{2}, 1\right)$, $(0, -1)$, $\left(-\dfrac{\sqrt{3}}{2}, 1\right)$, $\left(-\dfrac{\sqrt{3}}{2}, -1\right)$

$(0, 1)$, $\left(\dfrac{\sqrt{3}}{2}, -1\right)$; tangente vertical en $(1, 0)$, $(-1, 0)$.

7. *a*) Punto cuspidal en $(0, 0)$.

Páginas 127-128-129

1. *a*) $\mathbf{f}' = (1, 2I, I^2)$ sobre $[0, \infty\rangle$; *c*) $\mathbf{f}''(t) = (0, 2, 2t)$, $t \in [0, \infty\rangle$;

e) $(0, 0, 1)$; *g*) $\cos + I \operatorname{sen} + I^2 \cos + \frac{4}{3} I^3$ sobre $[0, \infty\rangle$;

i) $[\exp \circ (-\alpha I)]\left(1-\alpha I, 2I-\alpha I^2, I^2 - \dfrac{\alpha}{3} I^2\right)$ sobre $[0, \infty\rangle$;

k) $D(\mathbf{f} \circ \varphi) = -\alpha(I, 2I^2, I^3) \exp \circ (-\alpha I)$;

m) $\frac{d}{dt}\mathbf{g}(t^2) = 2t(-\text{sen}\ t^2, \cos t^2, 1)$;

o) $D(|\mathbf{f}|) = \frac{I + 2I^3 + \frac{1}{3}I^5}{I^{1/2} \circ (I^2 + I^4 + \frac{1}{9}I^6)}$.

2. *a*) $a\omega(-\text{sen}\ \omega t, \cos \omega t)$.

3. *a*) $\{t \mid t \in D_{\mathbf{f}}, \mathbf{f}(t) \neq 0\}$; $D_t|\mathbf{f}(t)| = \frac{\mathbf{f}(t)\cdot\mathbf{f}'(t)}{|\mathbf{f}(t)|}$.

6. *a*) $\text{Comp}_{\mathbf{f}(t)}\,\mathbf{f}'(t) = 0$; $\text{Comp}_{\mathbf{f}(t)}\,\mathbf{f}''(t) = -r\omega^2$.

7. $(-1, -\pi)$. **11.** *d*) 1) 2, 3) -3, 5) $-\frac{1}{2}$.

Página 131

1. *a*) $\Delta\mathbf{f}(0; 10^{-3}) = (10^{-3}, 10^{-6}, 10^{-9})$; $d\mathbf{f}(0; 10^{-3}) = (10^{-3}, 0, 0)$;

c) $\Delta\mathbf{f}(0; 10^3) = (10^3, 10^6, 10^9)$; $d\mathbf{f}(0; 10^3) = (10^3, 0, 0)$;

e) $\Delta\mathbf{f}(10^3; 10^{-1}) = (0.1, 200.01, 30030.001)$;

$d\mathbf{f}(10^3; 10^{-1}) = (0.1, 200, 30000)$.

2. *a*) $(1, 10^{-3}, 10^{-3})$; *c*) $(0.999, 0, 0.999)$.

Páginas 135-136

1. *a*) $(1, \frac{2}{3}, e-1)$; *c*) $\left(\frac{1}{2}\ln\frac{17}{5}, 2\sqrt{17} - \sqrt{5} + \frac{1}{2}\ln\frac{4+\sqrt{17}}{2+\sqrt{5}}, 240\right)$.

2. *a*) $\mathbf{x}(t) = \mathbf{c}t$; *c*) $\mathbf{x}(t) = (\cos \omega t, \text{sen}\ \omega t, 0)$.

3. $\mathbf{x}(t) = \mathbf{v}_0 t + \mathbf{x}_0$. **5.** 152 pies/seg.

7. $\mathbf{x}(t) = \mathbf{c}_1 \cos \omega t + \mathbf{c}_2\ \text{sen}\ \omega t$, $\omega^2 = \frac{k}{m}$.

8. *a*) 1) $\mathbf{x}(t) = \mathbf{0}$; 2) $\mathbf{x}(t) = \mathbf{x}_0 \cos \omega t$; 3) $\mathbf{x}(t) = \frac{1}{\omega}\mathbf{v}_0\ \text{sen}\ \omega t$;

c) $\mathbf{x}_0 \perp \mathbf{v}_0$ y $|\mathbf{x}_0| = r$, $|\mathbf{v}_0| = \omega r$.

Páginas 141-142-143

1. $\sqrt{2} + \ln(1+\sqrt{2})$. **3.** $8a$. **5.** $\frac{\sqrt{3}}{2} + \ln\frac{1+\sqrt{3}}{\sqrt{2}}$.

7. $\sqrt{2}a\ \text{senh}\ \frac{t}{a}$. **9.** 8. **13.** $\frac{d}{2}[\sqrt{2} - \ln(\sqrt{2}-1)]$.

Páginas 146-147-148-149

1. *a*) $\mathbf{T}(t) = \dfrac{1}{\sqrt{t^2+1}}(t, 1)$, $\mathbf{N}(t) = \dfrac{1}{\sqrt{t^2+1}}(1, -t)$;

   c) $\mathbf{T}(t) = \dfrac{1}{\sqrt{a^2 \operatorname{senh}^2 t + b^2 \cosh^2 t}}(a \operatorname{senh} t, b \cosh t)$;

   $\mathbf{N}(t) = \dfrac{1}{\sqrt{a^2 \operatorname{senh}^2 t + b^2 \cosh^2 t}}(b \cosh t, -a \operatorname{senh} t)$;

   e) $\mathbf{T} = \dfrac{\mathbf{a}}{|\mathbf{a}|}$, $\mathbf{N}$ no definida.

2. *a*) $\{s(1,1) \mid s \in \mathrm{R}\}$;

   c) ninguna recta tangente cuando $t = 0$; $\{(1,1,1)+s(2,3,4) \mid s \in \mathrm{R}\}$ cuando $t = 1$.

3. *a*) $\mathbf{v}(0) = \mathbf{v}(1) = 20\pi(1, 0)$, $|\mathbf{v}(0)| = |\mathbf{v}(1)| = 20\pi$;
   $\mathbf{a}(0) = \mathbf{a}(1) = 40\pi^2(0, -1)$, $\mathrm{Comp}_{\mathbf{T}(0)}\, \mathbf{a}(0) = \mathrm{Comp}_{\mathbf{T}(1)}\, \mathbf{a}(1) = 0$;
   $\mathrm{Comp}_{\mathbf{N}(0)}\, \mathbf{a}(0) = \mathrm{Comp}_{\mathbf{N}(1)}\, \mathbf{a}(1) = 40\pi^2$;

   c) $\mathbf{v}(0) = \mathbf{0}$, $|\mathbf{v}(0)| = 0$, $\mathbf{a}(0) = 2\pi(0,1)$;
   $\mathrm{Comp}_{\mathbf{T}(0)}\, \mathbf{a}(0) = 2\pi$, $\mathrm{Comp}_{\mathbf{N}(0)}\, \mathbf{a}(0) = 0$;
   $\mathbf{v}(1) = 2\pi(0, -1)$, $|\mathbf{v}(1)| = 2\pi$, $\mathbf{a}(1) = (4\pi^2, -2\pi)$,
   $\mathrm{Comp}_{\mathbf{T}(1)}\, \mathbf{a}(1) = 2\pi$, $\mathrm{Comp}_{\mathbf{N}(1)}\, \mathbf{a}(1) = -4\pi^2$;

   e) $\mathbf{v}(0) = \mathbf{v}(1) = \left(0{,}200\pi, \dfrac{1}{2\pi}\right)$, $|\mathbf{v}(0)| = |\mathbf{v}(1)| = \sqrt{40\,000\pi^2 + \dfrac{1}{4\pi^2}}$;
   $\mathbf{a}(0) = \mathbf{a}(1) = (-40\,000\pi^2, 0, 0)$, $\mathrm{Comp}_{\mathbf{T}(0)}\, \mathbf{a}(0) = \mathrm{Comp}_{\mathbf{T}(1)}\, \mathbf{a}(1) = 0$,
   $\mathrm{Comp}_{\mathbf{N}(0)}\, \mathbf{a}(0) = \mathrm{Comp}_{\mathbf{N}(1)}\, \mathbf{a}(1) = 40\,000\pi^2$.

5. *a*) $\mathbf{v}(0) = \mathbf{v}(1) = (20\pi, 0)$, $l'(0) = l'(1) = 20\pi$,
   $\mathbf{a}(0) = \mathbf{a}(1) = (0, -40\pi^2)$, $\mathrm{Comp}_{\mathbf{u}(0)}\, \mathbf{a}(0) = \mathrm{Comp}_{\mathbf{u}(1)}\, \mathbf{a}(1) = -40\pi^2$,
   $\mathrm{Comp}_{\mathbf{T}(0)}\, \mathbf{a}(0) = \mathrm{Comp}_{\mathbf{T}(1)}\, \mathbf{a}(1) = 0$,
   $\mathrm{Comp}_{\mathbf{N}(0)}\, \mathbf{a}(0) = \mathrm{Comp}_{\mathbf{N}(1)}\, \mathbf{a}(1) = 40\pi^2$.

   c) $\mathbf{v}(0) = \mathbf{0}$, $l'(0) = 0$, $\mathbf{a}(0) = 2\pi(0,1)$, $\mathrm{Comp}_{\mathbf{u}(0)}\, \mathbf{a}(0) = 0$,
   $\mathrm{Comp}_{\mathbf{T}(0)}\, \mathbf{a}(0) = 2\pi$, $\mathrm{Comp}_{\mathbf{N}(0)}\, \mathbf{a}(0) = 0$; $\mathbf{v}(1) = 2\pi(0, -1)$,
   $l'(1) = 2\pi$, $\mathbf{a}(1) = (4\pi^2, -2\pi)$, $\mathrm{Comp}_{\mathbf{u}(1)}\, \mathbf{a}(1) = -4\pi^2$,
   $\mathrm{Comp}_{\mathbf{T}(1)}\, \mathbf{a}(1) = 2\pi$, $\mathrm{Comp}_{\mathbf{N}(1)}\, \mathbf{a}(1) = 4\pi^2$;

   e) $\mathbf{v}(0) = \left(-1, \dfrac{\pi}{4}\right)$, $l'(0) = \dfrac{1}{4}\sqrt{16+\pi^2}$, $\mathbf{a}(0) = \left(1 - \dfrac{\pi^2}{16}, -\dfrac{\pi}{2}\right)$,

   $\mathrm{Comp}_{\mathbf{u}(0)}\, \mathbf{a}(0) = 1 - \dfrac{\pi^2}{16}$, $\mathrm{Comp}_{\mathbf{T}(0)}\, \mathbf{a}(0) = -\dfrac{1}{4}\sqrt{16+\pi^2}$,

$$\text{Comp}_{\mathbf{N}(0)}\,\mathbf{a}(0)=\frac{\pi}{16}\sqrt{16+\pi^2},\ \mathbf{v}(1)=\frac{1}{\sqrt{2}e}\left(-1-\frac{\pi}{4},\ -1+\frac{\pi}{4}\right),$$

$$l'(1)=\frac{1}{4e}\sqrt{16+\pi^2},\ \mathbf{a}(1)=\frac{1}{\sqrt{2}e}\left(1-\frac{\pi^2}{16}+\frac{\pi}{2},\ 1-\frac{\pi^2}{16}-\frac{\pi}{2}\right),$$

$$\text{Comp}_{\mathbf{u}(1)}\,\mathbf{a}(1)=\frac{1}{e}\left(1-\frac{\pi^2}{16}\right),\ \text{Comp}_{\mathbf{T}(1)}\,\mathbf{a}(1)=-\frac{1}{4e}\sqrt{16+\pi^2},$$

$$\text{Comp}_{\mathbf{N}(1)}\,\mathbf{a}(1)=\frac{\pi}{16e}\sqrt{16+\pi^2}.$$

7. $\mathbf{T}(0)=(1, 0, 0),\ \mathbf{N}(0)=(0, 1, 0),\ \mathbf{B}(0)=(0, 0, 1);\ z=0.$

8. *a*) $\mathbf{T}(0)=\frac{1}{\sqrt{2}}(1, 0, 1),\ \mathbf{N}(0)=(0, 1, 0),\ \mathbf{B}(0)=\frac{1}{\sqrt{2}}(-1, 0, 1);$

$x-z=0.$

Páginas 152-153

3. *a*) $\kappa=0$; *c*) $\kappa(\theta)=\dfrac{ab}{(a^2\operatorname{sen}^2\theta+b^2\cos^2\theta)^{3/2}}$; *e*) $\kappa(\theta)=\sqrt{\dfrac{\theta^4+5\theta^2+8}{(\theta^2+2)^3}}.$

5. *a*) $\kappa(\theta)=\dfrac{\theta^2+2}{|a|\,(\theta^2+1)^{3/2}}.$ **9.** $\dfrac{3}{9t^4+9t^2+1}.$ **11.** $\tau(0)=-1.$

Páginas 159-160

1. *a*) $\mathbf{v}(0)=\mathbf{v}(T)=\dfrac{20\pi}{T}(0, 1, 0),\ \mathbf{v}\left(\dfrac{T}{2}\right)=-\dfrac{20\pi}{T}(0, 1, 0),$

$$\mathbf{a}(0)=\mathbf{a}(T)=-\frac{40\pi^2}{T^2}(1, 0, 0),\ \mathbf{a}\left(\frac{T}{2}\right)=\frac{40\pi^2}{T^2}(1, 0, 0),\ a_{\mathbf{T}}=0,$$

$$a_{\mathbf{N}}=\frac{40\pi^2}{T^2},\ \omega=\left(0, 0, \frac{2\pi}{T}\right),\ \text{momento angular}=\left(0, 0, \frac{200\pi m}{T}\right);$$

c) $\mathbf{v}(0)=\mathbf{v}(T)=\dfrac{2\pi}{T}(a, 0, b),\ \mathbf{v}\left(\dfrac{T}{2}\right)=-\dfrac{2\pi}{T}(a, 0, b),$

$$\mathbf{a}(0)=\mathbf{a}(T)=-\frac{4\pi^2}{T^2}(0, a, 0),\ \mathbf{a}\left(\frac{T}{2}\right)=\frac{4\pi^2}{T^2}(0, a, 0),$$

$$a_{\mathrm{T}}(0) = a_{\mathrm{T}}\left(\frac{T}{2}\right) = a_{\mathrm{T}}(T) = 0,\ a_{\mathrm{N}}(0) = a_{\mathrm{N}}\left(\frac{T}{2}\right) = a_{\mathrm{N}}(T) = \frac{4\pi^2 a}{T^2},$$

$$\omega(0) = \omega\left(\frac{T}{2}\right) = \omega(T) = \frac{2\pi}{T}\left(\frac{b}{a}, 0, -1\right),$$

$$\text{momento angular} = \frac{2\pi m}{T}(ab, 0, -a^2).$$

Páginas 160-161

2. $P = (5\cos\theta - \cos 5\theta,\ 5\,\mathrm{sen}\,\theta - \mathrm{sen}\,5\theta)$. **5.** 24.

7. $\mathbf{f}(t) = \mathbf{a}e^{rt} + \mathbf{b}$ o $\mathbf{f}(t) = \mathbf{a}t + \mathbf{b}$.

8. $\mathbf{f}(t) = (500t,\ 1000t,\ -16t^2 + 1600)$, 11,284,

$$a_{\mathrm{T}} = \frac{1024t}{\sqrt{1024t^2 + 125 \times 10^4}},\quad a_{\mathrm{N}} = \sqrt{\frac{1280 \times 10^6}{1024t^2 + 125 \times 10^4}},\quad \rho(5) = 40950.$$

9. *a*) $\mathbf{v} = (0, 5)$, $\mathbf{a} = \left(-\frac{25}{\sqrt{2}}e^{-\pi/4}, 0\right)$; *b*) $\frac{1}{5}\sqrt{2}(e^{\pi} - 1)$;

c) $\mathbf{f}(t) = \left[\frac{5}{\sqrt{2}}t + 1\right]\left(\cos \ln\left(\frac{5}{\sqrt{2}}t + 1\right),\ \mathrm{sen}\,\ln\left(\frac{5}{\sqrt{2}}t + 1\right)\right)$.

10. *a*) $\mathbf{T}(\varphi) = (\cos\varphi,\ \mathrm{sen}\,\varphi)$, $\mathbf{N}(\varphi) = (-\mathrm{sen}\,\varphi,\ \cos\varphi)$,

$\kappa(\varphi) = \dfrac{1}{a\varphi}$, $L = 2\pi^2 a$;

b) $\mathbf{T}(\theta) = (\cos 3\theta,\ \mathrm{sen}\,3\theta)$, $\mathbf{N}(\theta) = (-\mathrm{sen}\,3\theta,\ \cos 3\theta)$,

$\kappa(\theta) = \dfrac{3}{10\,|\mathrm{sen}\,2\theta|}$, $L = 40$.

Páginas 166-167

1. *a*) Abierto; *c*) cerrado; *e*) cerrado.

Página 173

1. *a*) $\mathscr{D}_f = \mathrm{R}^3$, $f(x, y, z) = xy^2 + z$;

c) $\mathscr{D}_f = \{(x, y, z) \mid x + z \neq 0\}$, $f(x, y, z) = \dfrac{xyz}{x+z}$.

3. $\ln 2$.

4. *a*) $\mathscr{D}_f = \{x \mid z + \cos xy \geqslant 0\}$, $f(x, y, z) = \sqrt{z + \cos xy}$;
c) $\mathscr{D}_f = \mathrm{R}^3$, $f(x, y, z) = e^{xyz}$.
5. *a*) $f = I_1 I_2{}^2 + I_2$.

Páginas 183-184-185

3. *a*) 2; *c*) sen 2; *e*) ˙19; *g*) 0.
7. *a*) Ningún límite; *c*) ningún límite; *e*) ningún límite.
10. *a*) $\lim\limits_{x \to -1} \lim\limits_{y \to 3} (x^3 + xy^2) = \lim\limits_{y \to 3} \lim\limits_{x \to -1} (x^3 + xy^2) = \lim\limits_{(x,y) \to (-1,3)} (x^3 + xy^2) = -10$;
c) $\lim\limits_{x \to 0} \lim\limits_{y \to 0} f(x, y) = \lim\limits_{y \to 0} \lim\limits_{x \to 0} f(x, y) = \lim\limits_{(x,y) \to (0,0)} f(x, y) = 0$.
11. $\lim\limits_{x \to 0} \lim\limits_{y \to 0} f(x, y) = 1$, $\lim\limits_{y \to 0} \lim\limits_{x \to 0} f(x, y) = -1$, $\lim\limits_{(x,y) \to (0,0)} f(x, y)$ no existe.

Páginas 187-188

3. *a*) $\{(x, y, z) \mid x \neq y\}$; *c*) $\{(x, y) \mid (x+1)^2 + y^2 \neq 1\}$. **5.** No.

Páginas 194-195

1. *a*) $df(\mathbf{x}; \mathbf{h}) = h_1 + h_2$; $\mathbf{D}f(\mathbf{x}) = (1, 1)$;
c) $df(\mathbf{x}; \mathbf{h}) = 2xh_1 + 2yh_2$; $\mathbf{D}f(\mathbf{x}) = (2x, 2y)$.
7. *a*) $\mathbf{D}f(x, y) = (3, 2)$; *c*) $\mathbf{D}f(x, y) = (2x, 0)$;
e) $\mathbf{D}f(x, y) = \left(\dfrac{2x}{y}, -\dfrac{x^2}{y^2}\right)$.

Páginas 200-201

1. *a*) $2xy$; *c*) $\dfrac{1}{\sqrt{42}}(x - 5z + 4)$.
2. *a*) $\dfrac{yz-4}{(x+y)^2}$; *c*) $\dfrac{x}{x+y}$.
3. *a*) $\dfrac{1}{\sqrt{13}}(3x - 2y) \cos xy$.

Páginas 206-207

1. *a*) $D_1 f(x, y) = 2xy^2$, $D_2 f(x, y) = 2x^2 y + 1$;

c) $D_1 f(x, y, z) = 2x \dfrac{z-y-2}{(x^2+1)^2}$, $D_2 f(x, y, z) = \dfrac{1}{x^2+1}$,

$D_3 f(x, y, z) = \dfrac{x^2}{x^2+1}$;

e) $D_1 f(x, y) = -\dfrac{y}{x^2+y^2}$, $D_2 f(x, y) = \dfrac{x}{x^2+y^2}$;

g) $D_1 f(x, y) = \dfrac{x}{\sqrt{x^2+y^2}}$, $D_2 f(x, y) = \dfrac{y}{\sqrt{x^2+y^2}}$.

Páginas 211-212

1. a) $2(I_1, I_2, I_3)$; c) $(2I_1 I_2 I_3^2, I_1^2 I_3^2, 2I_1^2 I_2 I_3)$;
e) $e^x(\cos(x+y) - \operatorname{sen}(x+y), -\operatorname{sen}(x+y))$.

2. a) $dw = 2xy^3\,dx + 3(x^2y^2+1)\,dy$;
c) $dw = \dfrac{x\,dx + y\,dy + z\,dz}{\sqrt{x^2+y^2+z^2}}$.

3. a) $\dfrac{1}{\sqrt{10}(x-y)^2}(-2x^2+6xy+3y+x)$; c) $\dfrac{1}{\sqrt{3}z^2}(yz+xz-xy)$.

4. a) $\sqrt{286}$; c) 1.

7. $df((1,2);(\frac{1}{2},\frac{1}{4})) = \frac{29}{100}$; $\Delta f((1,2);(\frac{1}{2},\frac{1}{4})) = \frac{19}{65}$. **9.** $\frac{104}{3}ft^3$.

Página 217

1. a) $D_{1,1}f(x, y) = y^2e^{xy}$, $D_{2,2}f(x, y) = 2x + x^2e^{xy}$,
$D_{1,2}f(x, y) = D_{2,1}f(x, y) = 2y + xye^{xy} + e^{xy}$;

c) $D_{1,1}f(x, y) = \dfrac{2y}{x^3}\sec^2\dfrac{y}{x}\left(1 + \dfrac{y}{x}\tan\dfrac{y}{x}\right)$,

$D_{2,2}f(x, y) = \dfrac{2}{x^2}\sec^2\dfrac{y}{x}\tan\dfrac{y}{x}$,

$D_{1,2}f((x, y) = D_{2,1}f(x, y) = -\dfrac{1}{x^3}\sec^2\dfrac{y}{x}\left(x + 2y\tan\dfrac{y}{x}\right)$.

2. a) $D_{1,1}f(-3, 2) = -20$, $D_{2,2}f(-3, 2) = 108$,
$D_{1,2}f(-3, 2) = D_{2,1}f(-3, 2) = -45$;
c) $D_{1,1}f(\frac{1}{2}, 3, -1) = \frac{3}{2}$, $D_{2,2}f(\frac{1}{2}, 3, -1) = -\frac{1}{128}$,
$D_{3,3}f(\frac{1}{2}, 3, -1) = \frac{3}{128}$, $D_{1,2}f(\frac{1}{2}, 3, -1) = D_{2,1}f(\frac{1}{2}, 3, -1) = \frac{1}{16}$,
$D_{1,3}f(\frac{1}{2}, 3, -1) = D_{3,1}f(\frac{1}{2}, 3, -1) = \frac{3}{16}$,
$D_{2,3}f(\frac{1}{2}, 3, -1) = D_{3,2}f(\frac{1}{2}, 3, -1) = -\frac{1}{128}$.

3. *a*) $D_{1,1,1}f(x, y) = y^3 \operatorname{sen} xy$, $D_{2,2,2}f(x, y) = x^3 \operatorname{sen} xy$, $D_{1,1,2}f(x, y) = D_{1,2,1}f(x, y) = D_{2,1,1}f(x, y) = xy^2 \operatorname{sen} xy - 2y \cos xy$, $D_{1,2,2}f(x, y) = D_{2,1,2}f(x, y) = D_{2,2,1}f(x, y) = x^2 y \operatorname{sen} xy - 2x \cos xy$.

Páginas 221-222

1. *a*) $f(x, y) = x^2 - 3xy + 4y^2$, $(x, y) \in R^2$;

c) $f(x, y) = 8c_1 x^3 + 2xy^2 - 5y^3$, $(x, y) \in R^2$;

e) $f(x, y) = (x-1)y^2 - \frac{1}{2}(x-1)^2 y^2 + \dfrac{c_2^{\,2}}{5c_1^{\,5}}(x-1)^5 - \dfrac{c_2}{2c_1^{\,4}}(x-1)^4 y + \dfrac{1}{3c_1^{\,3}}(x-1)^3 y^2$, $x \geqslant 0$, $y \in R$;

g) $f(x, y) = 1 - \frac{1}{2}x^2 y^2 + \dfrac{1}{7!}[c_2^{\,7} x^7 + 7c_1 c_2^{\,6} x^6 y + 21(c_1^{\,2} c_2^{\,5} - 20c_2^{\,3})x^5 y^2 + 35(c_1^{\,3} c_2^{\,4} - 36c_1 c_2^{\,2})x^4 y^3 + 35(c_1^{\,4} c_2^{\,3} - 36c_1^{\,2} c_2)x^3 y^4 + 21(c_1^{\,5} c_2^{\,2} - 20c_1^{\,3})x^2 y^5 + 7c_1^{\,6} c_2 xy^6 + c_1^{\,7} y^7] \operatorname{sen} c_1 c_2 - \dfrac{1}{7!}[42c_2^{\,5} x^6 y + 210c_1 c_2^{\,4} x^5 y^2 - 420(c_1^{\,2} c_2^{\,3} - 2c_2)x^4 y^3 - 420(c_1^{\,3} c_2^{\,2} - 2c_1)x^3 y^4 - 210c_1^{\,4} c_2 x^2 y^5 + 42c_1^{\,5} xy^6] \cos c_1 c_2$;

i) $f(x, y) = \sqrt{3}[1 + \frac{3}{2}(x-1) + \frac{1}{6}(y-3) + \frac{3}{8}(x-1)^2 + \frac{1}{4}(x-1)(y-3) - \frac{1}{72}(y-3)^2 - \frac{1}{16}(x-1)^3 + \frac{1}{16}(x-1)^2(y-3) - \frac{1}{48}(x-1)(y-3)^2 + \frac{1}{432}(y-3)^3] + \dfrac{\sqrt{c_1^{\,3} c_2}}{128}\left[\dfrac{3}{c_1^{\,4}}(x-1)^4 - \dfrac{4}{c_1^{\,3} c_2} \times (x-1)^3 (y-3) - \dfrac{6}{c_1^{\,2} c_2^{\,2}}(x-1)^2 (y-3)^2 + \dfrac{12}{c_1 c_2^{\,3}}(x-1)(y-3)^3 - \dfrac{5}{c_2^{\,4}}(y-3)^4\right]$.

3. *a*) 0; *c*) 0.

Páginas 226-227

1. *a*) $2x + 3y + \sqrt{3}z = 16$; *c*) $6x + 9\sqrt{15}y = 72$.

3. *a*) $z = x + y$; *c*) $z = x - \frac{2}{3}y$.

Páginas 234-235

1. *c*) $z = \pm\sqrt{12 - 4x^2 - 3y^2}$ $\quad \dfrac{\partial z}{\partial x} = \dfrac{-4x}{\pm\sqrt{12 - 4x^2 - 3y^2}}$,

$$\frac{\partial z}{\partial y} = \frac{-3y}{\pm\sqrt{12-4x^2-3y^2}};$$

e) cualquier punto (x, y, z) sobre la superficie tal $z \neq 0$

$$\frac{\partial z}{\partial x} = -\frac{4x}{z}, \quad \frac{\partial z}{\partial y} = -\frac{3y}{z};$$

g) cualquier punto (x, y, z) sobre la superficie tal que $x \neq 0$

$$\frac{\partial x}{\partial y} = -\frac{3y}{4x}, \quad \frac{\partial x}{\partial z} = -\frac{z}{4x}.$$

3. $f(x, y) = \dfrac{-1+\sqrt{1-12xy^2}}{2y}$.

4. *a*) $\dfrac{\partial z}{\partial x} = \dfrac{\operatorname{sen} yz}{4-xy\cos yz}, \quad \dfrac{\partial z}{\partial y} = \dfrac{xz\cos yz}{4-xy\cos yz};$

c) $\dfrac{dz}{dx} = -\dfrac{x-4xz^5\sqrt{x^2+y^2+z^2}}{z-10x^2z^4\sqrt{x^2+y^2+z^2}} \quad \dfrac{\partial z}{\partial y} = -\dfrac{y}{z-10x^2z^4\sqrt{x^2+y^2+z^2}}$

5. *a*) $\dfrac{dy}{dx} = -\dfrac{y^3+3y}{3xy^2+3x}$; *c*) $\dfrac{dy}{dx} = -\dfrac{e^y}{xe^y+3}$; *e*) $\dfrac{dy}{dx} = \dfrac{1}{\cosh y+2y}$.

Páginas 244-245

1. *a*) $f(0, 0) = 0$ mín.; *c*) $f(-1, \frac{3}{2}) = \frac{3}{4}$ mín.

2. *a*) $(0, 0)$ punto de ensilladura; $(-2, 0)$ mín. rel.;

c) $(0, 0)$ punto de ensilladura;

e) puntos de ensilladura en $(m\pi, n\pi)$; mín. rel. en $\left((4m+1)\frac{\pi}{2}, (4n+3)\frac{\pi}{2}\right)$

y $\left((4m+3)\frac{\pi}{2}, (4n+1)\frac{\pi}{2}\right)$;

máx. rel. en $\left((4m+1)\frac{\pi}{2}, (4n+1)\frac{\pi}{2}\right)$ y $\left((4m+3)\frac{\pi}{2}, (4n+3)\frac{\pi}{2}\right)$;

g) $(-\frac{5}{2}, 0)$ mín. rel.; $(-2, -\frac{1}{2})$ y $(-2, \frac{1}{2})$ puntos de ensilladura.

3. 7, (3, 3, 4), (6, 5, 10).

6. *a*) $(-1, \frac{3}{2})$ mín. rel.; *c*) $(\frac{1}{2}, 0)$ mín. rel.

7. Cubo de lado $2\sqrt{2}$.

Páginas 246-247

3. No.

4. $\lim\limits_{x\to 0}\lim\limits_{y\to 0} f(x,y)=0$, $\lim\limits_{y\to 0}\lim\limits_{x\to 0} f(x,y)=1$; $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ no existe.

5. $D_1 f(x,y,z) = 2x \operatorname{sen} yz$, $D_2 f(x,y,z) = x^2 z \cos yz$,
$D_3 f(x,y,z) = x^2 y \cos yz$,

$$D_{\mathbf{u}} f(x,y,z) = \frac{1}{\sqrt{6}}(-2x \operatorname{sen} yz + 2x^2 z \cos yz + x^2 y \cos yz).$$

6. $D_1 f(0,0)=0$, $D_2 f(0,0)=0$. **7.** $-\dfrac{1}{\sqrt{37}}(0,1,6)$, $\sqrt{37}$.

10. $f(x,y) = y + xy + \frac{1}{2}x^2 y - \frac{1}{6}y^3 + R_3$.

11. $y = 3$. **12.** $z = x$. **13.** $f(8,4) = -11$ mín.

15. *a)* $\mathscr{E}_i = \mathscr{E}$, $\mathscr{E}_b = \{(0,0)\} \cup \{(x,y) \mid x^2+y^2 = 16\}$,
$\mathscr{E}_e = \{(x,y) \mid x^2+y^2 > 16\}$, $\overline{\mathscr{E}} = \{(x,y) \mid x^2+y^2 \leqslant 16\}$,
ningún punto aislado;
c) $\mathscr{E}_i = \mathscr{E}$, $\mathscr{E}_b = \{(x,y,z) \mid z = x^2+y^2\}$ $\mathscr{E}_e = \{(x,y,z) \mid z < x^2+y^2\}$,
$\overline{\mathscr{E}} = \{(x,y,z) \mid z \geqslant x^2+y^2\}$, ningún punto aislado.

Páginas 253-254

2. *a)* $(2,3,2)$; *c)* $(\operatorname{sen} 6, \tan \frac{2}{3})$.

Páginas 260-261-262

1. *a)* $\begin{pmatrix} 10 & 2 & 9 \\ 5 & 12 & -6 \\ -2 & 3 & 6 \end{pmatrix}$; *c)* $\begin{pmatrix} 4 & 9 & 2 \\ 8 & 23 & 22 \\ 11 & 8 & -5 \end{pmatrix}$; *e)* $\begin{pmatrix} -12 & 1 & -3 \\ 49 & 75 & 15 \\ 29 & 10 & 0 \end{pmatrix}$.

9. $\begin{pmatrix} 6 & 3 & 4 \\ 3 & 8 & 4 \\ 8 & -2 & 2 \end{pmatrix}$.

Páginas 267-268

1. *a)* $D\mathbf{f}(x,y,z) = \begin{pmatrix} z & 0 & x \\ 0 & 1 & 1 \end{pmatrix} d\mathbf{f}((x,y,z);(dx,dy,dz)) = (z\,dx + x\,dz, dy + dz)$;

c) $D\mathbf{f}(x,y,z)=\begin{pmatrix}\frac{1}{y} & \frac{-x}{y^2} & 0\\ 0 & 2 & 0\\ z^2 & 0 & 2xz\end{pmatrix}$

$$d\mathbf{f}((x,y,z);(dx,dy,dz))=\left(\frac{1}{y}dx-\frac{x}{y^2}dy,\,2\,dy,\,z^2\,dx+2xz\,dz\right).$$

Páginas 277-278-279

1. $\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial u},\ \frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial v},\ \frac{\partial z}{\partial w}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial w}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial w}.$

3. $\frac{\partial t}{\partial u}=3u^2w^3\operatorname{sen}v,\quad \frac{\partial t}{\partial v}=u^3w^3\cos v,\quad \frac{\partial t}{\partial w}=3u^3w^2\operatorname{sen}v.$

5. $D_1F(r,\theta,z)=\cos\theta D_1f(r\cos\theta,r\operatorname{sen}\theta,z)+\operatorname{sen}\theta D_2f(r\cos\theta,r\operatorname{sen}\theta,z)$
$D_2F(r,\theta,z)=-r\operatorname{sen}\theta D_1f(r\cos\theta,r\operatorname{sen}\theta,z)+r\cos\theta D_2f(r\cos\theta,r\operatorname{sen}\theta,z)$
$D_3F(r,\theta,z)=D_3f(r\cos\theta,r\operatorname{sen}\theta,z).$

6. a) $D_1F(u,v)=2u\cos^2 a,\ D_2F(u,v)=0.$

13. $\frac{\partial u}{\partial x}=\frac{4xyu^2v+yzv^2+3z}{yz^2-4x^2yuv},\quad \frac{\partial u}{\partial y}=\frac{4y^2v-xzv^2-z^2u}{yz^2-4x^2yuv}$

$\frac{\partial u}{\partial z}=\frac{2xv^2-zu}{z^2-4x^2uv},\quad \frac{\partial v}{\partial x}=\frac{2x^2yuv^2-6x^2u-2xyzu^2-yz^2v}{xyz^2-4x^3yuv}$

$\frac{\partial v}{\partial y}=\frac{2x^3uv^2+2x^2zu^2-2y^2z}{xyz^2-4x^3yuv},\quad \frac{\partial v}{\partial z}=\frac{2x^2yu^2-xyzv}{xyz^2-4x^3yuv}.$

Páginas 286-287

1. c) $y=z.$ 2. a) $x+y=10\sqrt{2}$; c) $x-y-\frac{\sqrt{2}}{4}z=0.$

5. a) $x=b\cos v\cos u,\ y=b\cos v\operatorname{sen}u,\ z=a\operatorname{sen}v,\ u\in[0,2\pi],$ $v\in[0,2\pi],\ a>b$;
c) $x=pv^2\cos u,\ y=pv^2\operatorname{sen}u,\ z=2pv,\ u\in[0,2\pi],\ v\in\mathrm{R}.$

Páginas 292-293

1. a) $F(0,0,\pm c)=c^2$;

c) $F\left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right)=\sqrt{3},\ F\left(-\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}\right)=-\sqrt{3}.$

3. Cubo de lado $2\sqrt{2}$.

7. $\left(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}, 0\right),\ \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0\right).$

Páginas 301-302-303

1. *a*) $\frac{3\pi}{2}$; *c*) 0; *e*) $\frac{11708}{105}$; *g*) $-\frac{704}{5}$. **2.** *a*) 63; *c*) $\frac{3\pi^2}{8}-1$.

3. *a*) $0, -2$. **4.** *a*) Sí; *c*) no. **5.** $4\pi^2$. **6.** *a*) 2π; *c*) 0.

Página 307

1. *a*) $U(\mathbf{x}) = -\frac{\mathbf{k}}{|\mathbf{x}|}$; *e*) $(0, 3, -2)$.

3. *a*) $U(\mathbf{x}) = \frac{1}{2}mk|\mathbf{x}|^2$;
c) esfera con centro en el origen; *e*) $(0, 0, \pm\sqrt{3})$.

4. *a*) No; *c*) 0. **5.** $\pm\frac{1}{\sqrt{17}}(6, 4, -17)$.

Páginas 308-309

1. *a*) $\begin{pmatrix} -\operatorname{sen} 2 & 0 \\ 0 & 2 \\ -2\operatorname{sen} 2 & 1 \end{pmatrix}$; *b*) $\begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} \\ -10 & 6 \end{pmatrix}$.

2. *a*) $D_1\mathbf{f}(x, y, z)=\left(\frac{z}{y}\sec^2\frac{x}{y}, yz\right),\ D_2\mathbf{f}(x, y, z)=\left(-\frac{xz}{y^2}\sec^2\frac{x}{y}, xz\right),$

$D_3\mathbf{f}(x, y, z)=\left(\tan\frac{x}{y}, xy\right),\ D_{1,1}\mathbf{f}(x, y, z)=\left(\frac{2z}{y^2}\sec^2\frac{x}{y}\tan\frac{x}{y}, 0\right),$

$D_{2,2}\mathbf{f}(x, y, z)=\left(\frac{2x^2z}{y^4}\sec^2\frac{x}{y}\tan\frac{x}{y}+\frac{2xz}{y^3}\sec^2\frac{x}{y}, 0\right),$

$D_{3,3}\mathbf{f}(x, y, z)=(0, 0),\ D_{1,2}\mathbf{f}(x, y, z)=$

$D_{2,1}\mathbf{f}(x, y, z)=\left(-\frac{2xz}{y^3}\sec^2\frac{x}{y}\tan\frac{x}{y}-\frac{z}{y^2}\sec^2\frac{x}{y}, z\right),$

$$D_{1,3}\mathbf{f}(x, y, z) = D_{3,1}\mathbf{f}(x, y, z) = \left(\frac{1}{y}\sec^2\frac{x}{y}, y\right),$$

$$D_{2,3}\mathbf{f}(x, y, z) = D_{3,2}\mathbf{f}(x, y, z) = \left(-\frac{x}{y^2}\sec^2\frac{x}{y}, x\right);$$

b) $D_1\mathbf{f}(x, y) = D_2\mathbf{f}(x, y) = (\frac{1}{2}(x+y)^{-1/2}, e^{x+y})\ D_{1,1}\mathbf{f}(x, y) =$
$D_{1,2}\mathbf{f}(x, y) = D_{2,1}\mathbf{f}(x, y) = D_{2,2}\mathbf{f}(x, y) = (-\frac{1}{4}(x+y)^{-3/2}, e^{x+y})$.

3. $\dfrac{\partial u}{\partial x} = -\dfrac{xu^2+y^2v}{2(x^2u+y^3)},\ \dfrac{\partial u}{\partial y} = \dfrac{2xyv-4y^2u-xyv}{2(x^2u+y^3)},$

$\dfrac{\partial v}{\partial x} = \dfrac{yu^2-xuv}{x^2u+y^3},\ \dfrac{\partial v}{\partial y} = -\dfrac{4xyu^2+x^2uv+2y^3v}{x^2yu+y^4}.$

4. $(y-1)D_1 f(xy-x, x^2+y)+2xD_2 f(xy-x, x^2+y)$. **6.** $4x+z+4 = 0$.

7. $\left(\frac{1}{2}, -\frac{1+\sqrt{2}}{2}\right), \left(-\frac{1}{2}, \frac{1+\sqrt{2}}{2}\right)$.

8. *a*) 24; *b*) 6π.

9. *a*) $U(\mathbf{x}) = \frac{1}{2}k \ln |\mathbf{x}|^2$; *b*) esfera con centro en el origen;

d) $\dfrac{k}{\sqrt{38}}$; *e*) $\frac{1}{2}k \ln \frac{38}{17}$; *f*) $(0, \pm 1, 0)$.

Página 321

1. *a*) 6; *c*) 12.

Páginas 340-341

4. *a*) 0.00025.

Página 347

a) $\frac{3625}{12}$; *c*) $\frac{7}{2}$; *e*) $\frac{1}{4}a^2\pi$; *g*) $\frac{1}{3}a^3$; *i*) $\frac{1}{2}$.

Página 352

1. *a*) 7; *c*) $\frac{45}{2}$; *e*) $\frac{1}{2}e^2-\frac{1}{2}e^{2/3}-\frac{2}{3}$.
2. *a*) 28; *c*) $\frac{5}{12}\pi\ \frac{5}{2}\pi-\frac{1}{2}\sqrt{3}$.

Página 357

a) 0; *c*) $\frac{5}{18}\sqrt{2}-\frac{2}{9}$; *e*) $\frac{2}{3}\pi$; *g*) $\frac{1}{2}$; *i*) 0; *k*) 20.

Páginas 363-364-365

1. *a)* $\frac{1}{5}(a, a)$; *c)* $\left(0, \dfrac{20}{3\pi}\right)$; *e)* $(\frac{10}{21}, \frac{10}{24})$.

2. *a)* $I_x = \frac{29}{420}$, $I_y = \frac{1}{20}$; *c)* $I_x = 32\pi - \frac{256}{5}$, $I_y = \frac{512}{5} + 32\pi$;
e) $I_x = \frac{1}{4}ab^3\pi$, $I_y = \frac{1}{4}a^3 b\pi$.

3. *a)* $A = \frac{1}{2}\pi + \frac{1}{3}$, $M_x = \frac{22}{15}$, $M_y = 0$, $I_x = \frac{4}{7} + \frac{1}{4}\pi$, $I_y = \frac{1}{4}\pi - \frac{2}{5}$,
$\bar{\mathbf{x}} = \left(0, \dfrac{44}{15\pi + 10}\right)$;
c) $A = \frac{1}{3}$, $M_x = M_y = \frac{3}{20}$, $I_x = I_y = \frac{3}{35}$, $\bar{\mathbf{x}} = (\frac{9}{20}, \frac{9}{20})$;
e) $A = \frac{9}{2}$, $M_x = \frac{81}{5}$, $M_y = \frac{9}{4}$, $I_x = \frac{8667}{140}$, $I_y = \frac{63}{20}$, $\bar{\mathbf{x}} = (\frac{1}{2}, \frac{18}{5})$.

4. *a)* $\frac{11}{60}$; *c)* $\frac{3}{20}\sqrt{2}$; *e)* $\frac{279}{70}$; *g)* $\frac{1313}{210}$.

11. $I_{\mathscr{L}} = \dfrac{1}{6(a^2+b^2)}\,[2a^2w^3h + 2b^2wh^3 +$
$6c^2wh + 3abw^2h^2 + 6acw^2h + 6bcwh^2]$.

Páginas 368-369

1. *a)* $\frac{5}{12}$; *c)* 2π. **3.** $\pi/2$.

Páginas 372-373

1. *a)* $\frac{3}{10}\pi$; *c)* $\frac{31}{30}\pi$. **3.** $\frac{23}{32}\pi$. **5.** $\frac{19}{3}\pi$.

Páginas 375-376

1. *a)* $\frac{1}{2}$; *c)* $\frac{1}{3}a^3$; *e)* $\frac{1}{18}\pi^2 + \ln 2$.

Páginas 381-382

a) $\frac{560}{3}$; *c)* $\frac{3}{4}(\pi - 2)$; *e)* $\frac{1}{6}abc$; *g)* $\frac{1}{24}abc^2$.

Páginas 385-386

1. *a)* 36; *c)* $\frac{3}{4}\cos 4 - \frac{3}{10}\cos 10 - \frac{9}{20}$; *e)* 1.

2. *a)* $\frac{1}{6}\pi a^3$; *c)* $\frac{5}{8}\pi$; *e)* $\frac{2}{45}a^5$.

Páginas 390-391

1. *a)* $\frac{4}{3}\pi abc$. **3.** $I_{yz} = \frac{1}{30}\pi a^3 bc$, $I_{zx} = \frac{1}{30}\pi ab^3 c$, $I_{xy} = \frac{1}{30}\pi abc^3$.

5. 12π. **7.** $\frac{2}{3}a^3$. **9.** $\frac{1}{3}\pi abc$.

Páginas 393-394

1. *a)* $(\frac{5}{2}, 4)$; *c)* $\left(3, \dfrac{44+12\pi}{16+3\pi}\right)$.

2. *a)* $I_b = \frac{580}{3}$, $I_c = \frac{100}{3}$; *c)* $I_b = 272+\frac{657}{8}\pi$, $I_c = 133.7$. **3.** $\frac{181}{16}\pi$.

Páginas 403-404

1. *a)* $2\sqrt{2}$; *b)* e^4-1; *c)* $\frac{1}{6}a^2$; *d)* $3\sqrt{3}a^2$.

2. *a)* $\dfrac{1}{e-1}(e-2, e/2)$; *b)* $(-\frac{1}{2}, \frac{2}{5})$; *c)* $(\frac{8}{5}, \frac{1}{2})$; *d)* $\left(\frac{1}{2}, \dfrac{7\pi}{60+5\pi}\right)$.

3. *a)* $\frac{243}{20}+\frac{9}{2}\sqrt{2}$; *b)* $\frac{729}{140}$. **4.** $2a^4-a^4 \ln 2$.

6. *a)* $\dfrac{15\,625}{3}$ lbs; *b)* $\frac{2003}{3}$ lbs; *c)* $\dfrac{10\,000}{3}$ lbs; *d)* $5\,000\pi$ lbs.

Páginas 425-426

1. *a)* $[\mathbf{abc}]$; *c)* $\dfrac{[\mathbf{abc}]}{60a_3b_3c_3}[(z_0+a_3+b_3+c_3)^5-(z_0+a_3+b_3)^5-$
$(z_0+a_3+c_3)^5-(z_0+b_3+c_3)^5+(z_0+a_3)^5+(z_0+b_3)^5+$
$(z_0+c_3)-{z_0}^5]$.

3. $\dfrac{45\,522}{35}$.

Páginas 441-442-443

1. $\frac{1}{9}$. **3.** $\frac{517}{2}$. **4.** *a)* $|a_1b_2-a_2b_1|$; *c)* 11. **5.** *c)* $\frac{2}{3}$.
6. *b)* $\frac{1}{8}\pi a^2$ senh 2. **7.** *c)* $8\pi a^3$; *e)* $\frac{32}{5}\pi a^5$.

Páginas 447-448

1. *a)* $\frac{3}{2}\pi a^2$; *c)* $12\pi-9\sqrt{3}$; *e)* $\sqrt{3}-\frac{1}{3}\pi$; *g)* $\frac{1}{32}\pi a^2$.

2. *a)* $(5a/6, 0)$; *c)* $\left(\dfrac{24\pi-33\sqrt{3}}{48\sqrt{3}-16\pi}, 0\right)$; *e)* $(\frac{1}{4}\pi, \frac{1}{4}\pi)$; *g)* $(\frac{1}{4}\pi a, 0)$.

3. *a)* $3a^4\pi/128$; *c)* $a^4(3\pi+8)/96$. **5.** $\pi/2$. **7.** $\frac{115}{96}$.
9. *a)* $\dfrac{\pi}{3}(3a-h)h^2$.

Página 449

1. $\frac{8}{3}\pi(2-\sqrt{3})$. **3.** $I_{xz} = I_{yz} = \frac{4}{15}\pi(16-9\sqrt{3})$. **5.** $\frac{4}{3}\pi a^2 b$.

Páginas 456-457

1. *a*) $1, 4, 9, 16, 25$; *c*) $-1, 1, -1, 1, -1$; *e*) $2, \frac{3}{2}, \frac{4}{3}, \frac{5}{4}, \frac{6}{5}$.
4. *a*) $(1, 0)$; *c*) diverge. **5.** *a*) 3.

Páginas 461-462

1. *a*) 2; *c*) 0; *e*) 0; *g*) 0. **2.** *a*) 1; *c*) 0.
4. *a*) 0; *c*) 0; *e*) 0. **10.** *a*) 1; *c*) 0; *e*) 0.

Páginas 466-467

1. *a*) ∞; *c*) $\frac{3}{5}$; *e*) $-\infty$; *g*) ∞. **3.** *a*) ∞.

Páginas 469-470

3. 2. **5.** *a*) $\sqrt{2}$; *c*) $-\sqrt{2}$.

Páginas 475-476

1. *a*) $-2, 2$; *c*) $-\infty, \infty$.
5. *a*) $\underline{\text{lím}}\, s_n = -\infty$, $\overline{\text{lím}}\, s_n = \infty$; *c*) $\underline{\text{lím}}\, s_n = -\frac{\sqrt{3}}{2}$, $\overline{\text{lím}}\, s_n = \frac{\sqrt{3}}{2}$;
e) $\underline{\text{lím}}\, s_n = 0$, $\overline{\text{lím}}\, s_n = \frac{3}{4}$.
11. $\underline{\text{lím}}\, s_n = a$, $\overline{\text{lím}}\, s_n = b$.

Páginas 480-481

3. *a*) máx. $= 8$, mín. $= -8$; *c*) máx. $= \frac{81}{4}$, mín. $= 0$;
e) máx. $= 5$, mín. $= -3$.

Páginas 487-488

1. *a*) $[-1, 1]$; *c*) $\langle -\infty, \infty \rangle$; *e*) $\langle -\infty, \infty \rangle$. **11.** Sí, no.

13. $f_n(x) = x, \; g_n(x) = x + \dfrac{1}{n}$.

Páginas 488-489

1. Converge para $|r| < 1$, diverge para $|r| \geqslant 1$.
2. *a*) $\frac{3}{4}$; *c*) $-\infty$; *e*) 0.

Página 496

1. *a*) Converge a $\frac{3}{4}$; *c*) diverge; *e*) converge a $\frac{2}{3}$.
2. *a*) Diverge **9.** No.

Páginas 506-507

1. *a*) Diverge; *c*) diverge; *e*) converge.
4. *a*) Converge; *c*) converge; *e*) converge.
6. *a*) Diverge; *c*) diverge; *e*) diverge; *g*) diverge; *i*) converge; *k*) diverge; *m*) converge.

Página 512

1. *a*) $n = 14$ (comparando con $\Sigma(\frac{2}{3})^k$);
c) $n = 10$ (comparando con $\Sigma(\frac{1}{2})^k$; *e*) $n = 15$; *g*) $n = 32$.

Páginas 516-517

1. *a*) $0.\overline{27}$; *c*) $0.\overline{692307}$. **3.** *a*) 1; *c*) $\frac{126}{990}$.

Páginas 520-521

1. *a*) $[-1, 1\rangle$; *c*) $\langle -1, 1\rangle$; *e*) $\langle -\infty, 0\rangle$; *g*) $\langle -\infty, \infty\rangle$.

Páginas 529-530

1. *a*) $1 + 0 - \dfrac{x^2}{2} + 0$; *c*) $1 + 0 + \dfrac{x^2}{2} + 0$;
e) $1 - (x-1) + (x-1)^2 - (x-1)^3$.

3. $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} (x-1)^k.$

5. $\sum_{k=0}^{\infty} 0x^k$, converge a $f(x)$ solamente para $x = 0$.

Páginas 536-537

2. *a*) $\langle -2, 0 \rangle$; *c*) $\{0\}$; *e*) $[1, 3]$.

3. *a*) $\sum_{k=0}^{\infty} \frac{(-1)^k}{5^{k+1}} x^k$; *c*) $\sum_{k=0}^{\infty} \left[\frac{(-1)^{k+1}}{3 \cdot 2^{k+1}} - \frac{1}{3}\right] x^k$;

e) $\sum_{k=0}^{\infty} \frac{1}{(2k)!} x^{2k}$; *g*) $\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{4k+2}$.

4. *a*) $\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!\,(2k+1)} x^{2k+1}$; *c*) $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{(2k)!\,2k} x^{2k}$. **7.** *b*) $\frac{3}{2}$.

Páginas 542-543

1. *a*) $x+\frac{1}{3}x^3+\frac{3}{10}x^5+\frac{16}{63}x^7$; *c*) $1+x-\frac{1}{3}x^3-\frac{1}{6}x^4$.

2. *a*) $1 + \sum_{k=1}^{\infty} \frac{(-1)^k 2^{2k-1}}{(2k)!} x^{2k}$. **4.** *a*) $1+\frac{1}{2}x^2+\frac{5}{24}x^4+\frac{61}{70}x^6$.

Páginas 543-544

1. *a*) Converge; *b*) converge; *c*) diverge; *d*) diverge.

2. 1.20. **3.** $[-2, 4\rangle$.

Páginas 552-553

1. *a*) $\frac{7}{2}$; *c*) $(4-\pi)/32$; *e*) diverge; *g*) $\pi/4$; *i*) $2\sqrt{3}\pi/9$; *k*) $\pi/2$.

2. *a*) -1; *c*) 2. **3.** *a*) 1; *c*) $\frac{3}{2}$.

Páginas 561-562-563

1. *a*) Diverge; *c*) converge;

e) diverge a causa del comportamiento en $x = -2$; *g*) converge

3. *a*) Converge; *c*) converge.

5. *a*) Converge; *c*) converge.

7. *a*) Converge; *c*) converge.
8. *a*) Converge; *c*) converge. **9.** $3a^2\pi/4$.

Páginas 565-566-567

1. *a*) $-2\int_0^{\pi}(1-y\,\text{sen}\,x)\,\text{sen}\,x\,dx$; *c*) $\int_{\pi/2}^{\pi}\cos(xy)\,dx$;

e) $\frac{1}{y}(e^{y^2}-e^{-y^2})+\int_{-y}^{y}e^{-xy}\,dx$; *g*) $\frac{\pi}{2}y-2\int_0^{y^2}\frac{xy}{y^4+x^2}\,dx$.

7. $y(x)=2x+3$. **9.** $y(x)=ax$.

Páginas 574-575-576

3. $F(y)=y^{-1}\ln^{-2}y$. **5.** $F(y)=\arctan y$. **9.** $\ln\frac{b}{a}$.

Página 580

1. *a*) Comparando con x^{-2}, tómese $b>\varepsilon^{-1}$;
c) comparando con $x^{-3/2}$, tómese $b>4\varepsilon^{-2}$;
e) comparando con x^{-3}, tómese $b>(2\varepsilon)^{-1/2}$;
g) comparando con e^{-x}, tómese $b>$ máx $\{1, -\ln\varepsilon\}$.
2. *a*) Comparando con $x^{-1/2}$, tómese $\delta<\frac{1}{4}\varepsilon^2$;
c) comparando con $x^{-1/2}$, tómese $\delta<\frac{1}{4}\varepsilon^2$.
3. *a*) $b=n\pi>(\pi/\varepsilon)^{1/2}$; *c*) $b=n\pi>e^{\pi/\varepsilon}$.

Páginas 581-582-583

1. *a*) Diverge; *b*) $\frac{1}{2}$; *c*) $\frac{1}{2}$; *d*) $\frac{1}{2}\,\text{arccos}\,2/5$; *e*) $\pi/(2ab)$; *f*) 2.
2. π. **3.** $8\sqrt{2}$.
6. *a*) Converge; *b*) converge; *c*) diverge; *d*) converge;
e) diverge (*Sugerencia:* véase problema 2, pág. 561); *f*) converge.

7. *a*) $-2y\int_0^{\pi}\text{sen}\,(xy^2)\,dx$; *b*) $-\int_3^7\cos(x-y)\,dx$;

c) $-2y\int_0^{\pi/y^2}\text{sen}\,(xy^2)\,dx+2\pi/y^3$;

d) $\int_x^{x^2}[-y(y-x)^4-4(y-x)^3]e^{-xy}\,dy+2xe^{-x^3}(x^2-x)^4$.

8. $-y^{-2}+y^{-2}\cos y^2 + \operatorname{sen} y^2$.

9. *c*) $6!\left(-\operatorname{sen} y + y - \dfrac{1}{3!}y^3 + \dfrac{1}{5!}y^5\right)$.

Páginas 593-594

1. *a*) $I + \operatorname{sen} + c$; *c*) $c + \displaystyle\int_0^t e^{-s^2}\,ds$.

2. *a*) $\ln x$; *c*) $10 + \displaystyle\int_0^x \frac{\operatorname{sen} t}{t}\,dt$.

3. *a*) $x+c$; *c*) ninguna solución.

5. *a*) ce^{2x}; *c*) $\frac{1}{4}e^x + ce^{-3x}$; *e*) $ce^{-x}+\frac{1}{2}(\operatorname{sen} x - \cos x)$; *g*) $10t + 3e^{-3t} - 2$.

7. *a*) $16t^2+28t+19$; *c*) $2t - \operatorname{sen} t$.

9. *a*) $\sqrt{x^2+c}$; *c*) $\sqrt{1-x^2}$.

11. *a*) Solamente si $a = 0$; *c*) no.

Páginas 598-599-600-601

1. ce^{-3t}. **3.** ce^{-bt}. **5.** $ce^{-3t}+\frac{1}{6}e^{3t}$.

7. $ce^{-bt} + \dfrac{1}{a+b}e^{at}$.

9. $-x\ln|x| + cx$. **11.** $\frac{1}{3}t^2$. **13.** $1-e^{(1-t^2)/2}$.

15. $e^{\frac{1}{2}t^2}\displaystyle\int_0^t e^{-\frac{1}{2}s^2}\,ds$. **17.** $e^{-\frac{1}{2}t^2}\left[2 + \displaystyle\int_0^t e^{\frac{1}{2}s^2}\,ds\right]$.

19. $x(t) = c_1e^t$, $y(t) = (c_2 + c_1t)e^t$.

21. $x(t) = (c_1 + c_2t)e^{\lambda_1 t}$, $y(t) = c_2e^{\lambda_1 t}$. **23.** 202.6 mín.

25. Torio B 93.67%, Torio C 4.56%. **27.** 24.5 mín.

29. $\frac{1}{4}e^{-\frac{1}{2}x^2}\left[\displaystyle\int_0^x e^{\frac{1}{4}t^2}\,dt + c\right]^2$. **31.** *a*) 115.5; *c*) 136.

Páginas 606-607

2. *a*) $2\sqrt{2}\operatorname{sen}(\theta + 3\pi/4)$.

Páginas 616-617-618

1. *a*) $x_1(t) = c_1e^{-3t}$, $x_2(t) = c_2e^{2t}$;

c) $x_1(t) = c_1 e^{-2t}$, $x_2(t) = (c_1 t + c_2)e^{-2t}$;
e) $x(t) = 10e^{2t} - 9e^{3t}$, $y(t) = 5e^{2t}$;
g) $x_1(t) = ae^{-3t}$, $x_2(t) = (2at + b)e^{-3t}$;
i) $x_1(t) = c_1 e^{\lambda_1 t}$, $x_2(t) = \dfrac{a_{21} c_1}{\lambda_1 - \lambda_2} e_{\lambda_1 t} + c_2 e^{\lambda_2 t}$;
k) $x_2(t) = c_2 e^{\lambda_2 t}$, $x_1(t) = (c_1 + a_{12} c_2 t)e^{\lambda_1 t}$ si $\lambda_1 = \lambda_2$ y
$x_1(t) = c_1 e^{\lambda_1 t} + \dfrac{a_{12} c_2}{\lambda_2 - \lambda_1} e^{\lambda_2 t}$ si $\lambda_1 \neq \lambda_2$.

2. *a)* $\lambda_1 = 2$, $\lambda_2 \equiv -3$, $\mathbf{v}^1 = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$, $\mathbf{v}^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$;
c) λ_1, λ_2 $\mathbf{v}^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{v}^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$;
e) λ_0 es un valor característico doble y todos los vectores característicos son de la forma $\begin{pmatrix} 0 \\ y_0 \end{pmatrix}$, $y_0 \neq 0$;
g) $\lambda_1 = 1$, $\lambda_2 = -b$, $\mathbf{v}^1 = \begin{pmatrix} 1+b \\ -c \end{pmatrix}$, $\mathbf{v}^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

4. *a)* $x_1(t) = c_1 e^t + 3c_2 e^{2t}$, $x_2(t) = c_2 e^{2t}$;
c) $x(t) = e^{2t}[c_1 \cos 2t + c_2 \operatorname{sen} 2t]$, $y(t) = e^{2t}[(-c_1 - 2c_2)\cos 2t + (2c_1 - c_2)\operatorname{sen} 2t]$; *e)* $x(t) = (2c_2 t + c_1)e^{3t}$, $y(t) = c_2 e^{3t}$;
g) $x(t) = \cos\sqrt{2}t + \sqrt{2}\operatorname{sen}\sqrt{2}t$, $y(t) = \sqrt{2}\operatorname{sen}\sqrt{2}t$;
i) $x(t) = [x_0 + (y_0 + ax_0)te^{-at}$, $y(t) = [y_0 - (y_0 + ax_0)at]e^{-at}$;
k) $x(t) = te^{-2t}$, $y(t) = \frac{3}{2}t^2 e^{-2t}$.

5. *b)* $x(\infty) = \dfrac{bM}{a+b}$, $y(\infty) = \dfrac{aM}{a+b}$;
c) contiene el punto $(x(\infty), y(\infty))$.

7. *a)* $x(t) = c_1 e^{-t} + c_3 e^{2t}$, $y(t) = c_2 e^{-t} + c_3 e^{2t}$, $z(t) = -(c_1 + c_2)e^{-t} + c_3 e^{2t}$.

Páginas 622-623-624

1. *a)* $c_1 e^t + c_2 e^{-t}$; *c)* $e^t(a_1 \cos\sqrt{2}t + b_1 \operatorname{sen}\sqrt{2}t)$; *e)* $c_1 + c_2 e^{-x/3}$.
2. *a)* $5 \operatorname{senh} 2t$; *c)* $x_0 + \frac{1}{3}\dot{x}_0 - \frac{1}{3}\dot{x}_0 e^{-3t}$; *e)* $e^{-2t}(2\cos t + 4\operatorname{sen} t)$.
8. *a)* $c_1 t^2 + c_2 t$; *c)* $c_1 + c_2 \ln t$.
9. *a)* $c_1 e^{-t} + c_2 e^{3t/2}$; *c)* $10^{-8}\cos 10^7 t$.

Páginas 630-631

4. *a)* $x(t) = \frac{1}{4}e^t - \frac{1}{4}(1+2t)e^{-t}$, $y(t) = \frac{1}{4}e^t + \frac{1}{4}(2t-1)e^{-t}$;
c) $x(t) = -e^{2t} + 2 + \frac{1}{2}t^2 + \frac{1}{6}t^3$, $y(t) = e^{2t} + 2 + \frac{1}{6}t^3$.

Páginas 634-635-636-637

1. *a)* $\frac{1}{2}e^{-2t}-3e^{-t}+\frac{5}{2}-2t+\frac{1}{2}t^2$; *c)* $-2e^{-t/2}-\frac{1}{6}e^{-3t}+\frac{5}{3}+\frac{1}{2}e^{-t}$;

e) $x_0 \cos ct + \frac{1}{c}\dot{x}_0 \operatorname{sen} ct + \frac{1}{1+c^2}\left(e^t - \cos ct - \frac{1}{c}\operatorname{sen} ct\right)$;

g) 0 para $t<0$, $\frac{1}{2}(e^{-t}\cos t + t - 1)$ para $0 \leqslant t < 1$,
$\frac{1}{2}[e^{-t}\cos t - e^{1-t}\cos(t-1)+1]$ para $1 \leqslant t$.

4. *a)* 0 para $t<0$, $1-\cos t$ para $0 \leqslant t < 1$, $1-\cos 1$ para $1 \leqslant t$;
c) 0 para $t<0$, $1-\cos t$ para $0 \leqslant t$.

6. *a)* 0 para $t<0$, $-e^{-t}+\frac{1}{2}e^{-2t}+\frac{1}{2}$ para $0 \leqslant t < 1$,
$e^{1-t}-e^{-t}+\frac{1}{2}(e^{-2t}-e^{2(1-t)})$ para $1 \leqslant t$;
c) 0 para $t<0$, $-e^{-t}+\frac{1}{2}e^{-2t}+\frac{1}{2}$ para $0 \leqslant t$.

Páginas 642-643

1. *a)* $x(t) = y(t)-y(t-1)$ donde $y(t) = 0$ para $t<0$ y
$y(t) = 1+\frac{1}{99}(e^{-10t}-100e^{-t/10})$ para $0 \leqslant t$, $\dot{x}(t) = \dot{y}(t)-\dot{y}(t-1)$
donde $\dot{y}(t) = 0$ para $t<0$ y $\dot{y}(t) = \frac{1}{9.9}(e^{-t/10}-e^{-10t})$ para $0 \leqslant t$.

2. *a)* $x(t) = y(t)-y(t-\frac{1}{5})$ donde $y(0) = 0$ para $t<0$ y
$y(t) = \frac{1}{5}-\frac{1}{4}e^{-t}+\frac{1}{20}e^{-5t}$ para $0 \leqslant t$.

3. *a)* $x(0) = 0$ para $t<0$ y $x(t) = \frac{1}{288}(9e^{-4t}-4e^{-6t}+12t-5)$ para $0 \leqslant t$;
c) $x(t) = y(t)-2y(t-1)+y(t-2)$ donde $y(t)$ es la respuesta del problema $3a$.

7. *a)* $x(t) = \frac{1}{39}(8\cos 3t + 12 \operatorname{sen} 3t)$, $y(t) = \frac{1}{39}(-20\cos 3t + 48 \operatorname{sen} 3t)$;
c) $x(t) = \frac{1}{6}(\cos 2t - 2\operatorname{sen} 2t) - \frac{1}{11}\operatorname{sen} 3t$,
$y(t) = \frac{1}{6}\cos 2t + \frac{1}{11}(-3\cos 3t + \operatorname{sen} 3t)$;
e) $\frac{1}{85}(6\cos t + 7\operatorname{sen} t) + \frac{1}{765}(6\cos 3t - 7\operatorname{sen} 3t)$.

Páginas 653-654

1. *a)* Punto de ensilladura; *c)* nodo estable;
e) foco inestable; *g)* foco estable.

4. *a)* $\frac{1}{\omega^4 - 41\omega^2 + 625}[(25-\omega^2)\cos \omega t + 3\omega \operatorname{sen} \omega t]$;
c) $x(t) = \frac{1}{2}(\cos t + \operatorname{sen} t) = -y(t)$.

5. *a)* $\frac{1}{5}$ para $\omega = 0$;
c) $10^3[\omega_0{}^2 + 10^8(16-\omega_0{}^2)^2]^{-1/2} \approx \frac{1}{4}\times 10^3$ para
$\omega_0 = 5\times 10^{-5}[64\times 10^8 - 2]^{1/2} \approx 4$.

Páginas 665-666-667

2. *a*) $v \geqslant -3$.
3. *a*) $[91^2 \times 10^6 + 225]^{-1/2} [91 \times 10^3 \cos 3t + 15 \operatorname{sen} 3t]$;
c) $a[(36-\omega^2)^2 + 16^2\omega^2]^{-1} [(36-\omega^2)\cos \omega t + 16\omega \operatorname{sen} \omega t]$;
e) $-\frac{35}{6} \cos 10t + \dfrac{1}{640{,}000.81} [0.225 \cos 30t + 200 \operatorname{sen} 30t]$;
g) $[\omega^2(\omega^2-14)^2 + (8-7\omega^2)^2]^{-1} \times [\omega(\omega^2-14)\cos \omega t + (8-7\omega^2) \operatorname{sen} \omega t]$.
4. *a*) $(c_1 + c_2 t)e^{-2t} - \frac{1}{50}(4 \cos 3t - 3 \operatorname{sen} 3t)e^{-t}$;
c) $c_1 e^{-5t} + c_2 e^{2t} + a[(\omega^2+10)^2 + 9\omega^2]^{-1} [3\omega \operatorname{sen} \omega t - (\omega^2+10) \times \cos \omega t]$.
5. *a*) $2.32 \times 10^{-2} E$; *c*) $10^{-5} E$.
7. $E^2 = \dfrac{1}{4\alpha}$, tómese α tan grande como sea posible.

Páginas 674-675

1. *a*) $x^2 + xy + y^2 = c$; *c*) $\frac{1}{2}x^2 - \dfrac{y}{x} = c$;
e) $x^3y^2 + 3x \operatorname{sen} y = c$; *g*) $y = c + \sqrt{a^2 - x^2}$.
3. *a*) $x + y = cy^2$; *c*) $e^x(xy^2 + y^4) = c$.
5. *a*) $xy = c_1$; *c*) $x^2 + (y-c)^2 = c^2$; *e*) $x^2 + y^2 - \ln x^2 = c$.

Páginas 682-683

1. *a*) $\frac{1}{3}$; *c*) **0**. **2.** *a*) 1; *c*) 1. **3.** $-\frac{1}{2}$.
4. *a*) $-\frac{1}{12}$; *c*) $\frac{1}{6}$. **5.** *a*) 1; *c*) **0**. **9.** **0**.
10. *a*) $-x^3y^2 + 3xy^2$; *c*) no conservativo.

Páginas 699-700-701

7. 1.90.

Páginas 706-707

1. *a*) $\frac{3}{4}e^{-2t} + \frac{3}{2}t^2 - \frac{3}{2}t + \frac{1}{4}$; *c*) $e^{-\frac{1}{4}t^2}$.
2. *c*) $x_2(t) = 1 + t + t^2 + t^3 + \frac{2}{3}t^4 + \frac{1}{3}t^5 + \frac{1}{9}t^6 + \frac{1}{63}t^7$;
c) $x_3(t) = t - t^2 + \frac{4}{3}t^3 - t^4 + \frac{2}{3}t^5 - \frac{2}{9}t^6 + \frac{4}{63}t^7$,
$y_3(t) = 1 - t + t^2 - \frac{2}{3}t^3 + \frac{1}{6}t^4$.

Páginas 715-716

1. $\pi + t + \dfrac{t^3}{3!} + \dfrac{t^4}{4!} + \cdots$

4. *a)* $y(x) = -2 + \frac{1}{2}(x-1) + \frac{3}{16}(x-1)^2 + \frac{7}{64}(x-1)^3 + \ldots;$
c) $x(t) = t + \frac{1}{2}t^2 + \frac{1}{3}t^3 + \ldots,\ y(t) = 1 + t - \frac{1}{6}t^3 - \ldots;$
e) $x(t) = 2 + 3t + \frac{7}{2}t^2 + 3t^3 + \ldots,\ y(t) = 2 + 6t + 9t^2 + \frac{32}{3}t^3 + \ldots$

5. *a)* $1 - 2k + x.$

Página 719

2. *a)* $x_n = 1,\ 1.1,\ 1.245,\ 1.427,\ 1.649,\ 1.915;$
c) $y_n = 0,\ 0.1,\ 0.222,\ 0.362,\ 0.521,\ 0.702;$
e) $x_n = 1,\ 1.4,\ 1.88,\ 2.18,\ 2.22,\ 2.22,\ y_n = 2,\ 1.8,\ 1.37,\ 0.82,\ 0.48,\ 0.32.$

5. *a)* $x_n = 0,\ 0.1,\ 0.199,\ 0.285,\ 0.380,\ 0.470,\ 0.556,\ 0.636,\ 0.710,\ 0.777,\ 0.836;$
c) $y_n = 0,\ 0.1,\ 0.222,\ 0.363,\ 0.523,\ 0.705.$

Páginas 734-735-736-737

9. *a)* $\dfrac{\pi^2}{3} + 4\sum_{n=1}^{\infty} \dfrac{(-1)^n \cos nx}{n^2};$

c) $2 \operatorname{senh} 1 \left[\frac{1}{2} + \sum_{n=1}^{\infty} \dfrac{(-1)^n}{1 + n^2\pi^2} (\cos n\pi x - n\pi \operatorname{sen} n\pi x) \right];$

e) $\dfrac{\pi}{2} - \dfrac{4}{\pi} \sum_{n=1}^{\infty} \dfrac{1}{(2n-1)^2} \cos(2n-1)x;$

g) $\frac{1}{4} - \dfrac{2}{\pi^2} \sum_{n=1}^{\infty} \dfrac{1}{(2n-1)^2} \times \cos(2n-1)x - \dfrac{1}{\pi} \sum_{n=1}^{\infty} \dfrac{(-1)^n \operatorname{sen} n\pi x}{n}.$

11. *a)* $\dfrac{4}{\pi} \sum_{n=1}^{\infty} \dfrac{\operatorname{sen} nx}{2n-1},\ 1;$ *c)* $\operatorname{sen} x,\ \dfrac{2}{\pi} - \dfrac{4}{\pi} \sum_{n=1}^{\infty} \dfrac{\cos nx}{4n^2-1};$

e) $\dfrac{8}{\pi} \sum_{n=1}^{\infty} \dfrac{\operatorname{sen}(2n-1)x}{(2n-1)^3},\ \dfrac{\pi^2}{6} - \sum_{n=1}^{\infty} \dfrac{\cos 2nx}{n^2}.$

Páginas 746-747

4. *a)* $s_3(x) = \frac{1}{4}e(-3 + 105x + 15x^2 - 175x^3) + \frac{1}{4}e^{-1}(33 - 765x - 105x^2 + 1295x^3),\ E_3{}^2 = -90e^2 + 1331 - 4921\,e^{-2} \approx 0;$
c) $s_3(x) = \frac{1}{4}\pi + \frac{5}{4}(3-\pi) \times (3x^2 - 1),\ E_3{}^2 = -22 + \frac{61}{4}\pi - \frac{21}{8}\pi^2 \approx 0.001.$

5. *a)* 2.01, 1.575, 1.335, 1.179; *c)* 1.439, 0.706, 0.434, 0.300.

Bibliografía

(1) Haaser, N. B., LaSalle, J. P., y Sullivan, J. A. *A Course in Mathematical Analysis, vol. I, Introduction to Analysis.* Ginn and Company, Boston, 1959.
(2) Apostol, T. M. *Mathematical Analysis.* Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1957.
(3) Apostol, T. M. *Calculus,* vol. 1. Blaisdell Publishing Company, Nueva York, 1961.
(4) Artin, E. *Calculus and Analytic Geometry.* The Mathematical Association of America, Buffalo, Nueva York, 1957.
(5) Buck, R. C. *Advanced Calculus.* McGraw-Hill Book Company, Inc., Nueva York, 1956.
(6) Courant, R. *Differential and Integral Calculus.* Traducida al inglés por McShane E. J. Interscience Publishers, Inc., Nueva York, vol. I, segunda edición, 1937, vol. II, 1936.
(7) Dieudonné, J. *Foundations of Modern Analysis.* Academic Press, Inc., Nueva York, 1960.
(8) Goursat, E. *Cours d'Analise.* Quinta edición, vol. I, Gauthier-Villars, París, 1956. La primera edición francesa del vol. I, ha sido traducida por Hedrick, E. R., como *A Course in Mathematical Analysis.* Ginn and Company, Boston, 1904. Reimpresa por Dover Publications, Inc., Nueva York.
(9) Kaplan, W. *Advanced Calculus.* Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1952.
(10) Khinchin, A. I. *A Course of Mathematical Analysis.* Traducido de la tercera edición rusa, 1957, Hindustan Publishing Corp. (India), Delhi, 1960.
(11) McShane, E. J. y Botts, T. *Real Analysis.* D. Van Nostrand Company, Inc., Princeton, Nueva Jersey, 1959.
(12) Nevanlinna, T. y R. *Absolute Analysis.* (Alemán) Springer-Verlag, Berlín, 1959.
(13) Nickerson, H. K., Spencer, D. C., y Steenrod, N. E. *Advanced Calculus.* D. Van Nostrand Company, Inc., Princeton, Nueva Jersey, 1959.
(14) Olmsted, J. M. H. *Advanced Calculus.* Appleton-Century-Crofts, Inc., Nueva York, 1961.
(15) Olmsted, J. M. H. *Real Variables.* Appleton-Century-Crofts, Inc., Nueva York, 1959.
(16) Ostrowski, A. *Vorlesungen über Differential-und Integralrechnung.* Verlag Birkhauser, Basel. Erster Band, 1945. Zweiter Band, 1951. Dritter Band, 1954.
(17) Poussin, Ch.-J. de la Vallee. *Cours d'Analyse Infinatesemale.* Tomo I, sexta edición; tomo II, séptima edición, Dover Publications, Nueva York, 1946.
(18) Rudin, W. *Principles of Mathematical Analysis.* McGraw-Hill Book Company, Inc., Nueva York, 1953.
(19) Taylor, A. E. *Advanced Calculus.* Ginn and Company, Boston, 1955.
(20) Whittaker, E. T. y Watson, G. N. *A Course in Modern Analysis,* cuarta edición. Cambridge University Press, Cambridge, 1940.

ALGEBRA LINEAL, VECTORES Y GEOMETRÍA ANALÍTICA

(21) Bellman, R. *Introduction to Matrix Analysis.* McGraw-Hill Book Company, Inc., Nueva York, 1960.
(22) Birkhoff, G. y MacLane, S. *A Survey of Modern Algebra,* edición revisada, The Macmillan Company, Nueva York, 1953.
(23) Faddeeva, V. N. *Computational Methods of Linear Algebra.* Traducido del ruso al inglés por Benster, C. D., Dover Publications, Inc., Nueva York, 1959.
(24) Gantmacher, F. R. *The Theory of Matrices.* Dos volúmenes traducidos del ruso al inglés por Hirsch, K. A., Chelsea Publishing Company, Nueva York, 1960.

(25) Gel'fand, I. M. *Lectures on Linear Algebra.* Traducido al inglés de la segunda edición rusa por Shenitzer, A., Interscience Publishers, Inc., Nueva York, 1961.
(26) Jaeger, A. *Introduction to Analytic Geometry and Linear Algebra.* Holt, Rinehart and Winston, Inc., Nueva York, 1960.
(27) Halmos, P. R. *Finite Dimensional Vector Spaces*, segunda edición. D. Van Nostrand Company, Inc., Princeton, Nueva Jersey, 1958.
(28) MacDuffee, C. C. *Vectors and Matrices.* Carus Matnematical Monographs, No. 7. The Mathematical Association of America, Buffalo, Nueva York, 1943.
(29) Murnaghan, F. D. *Analytic Geometry.* Prentice-Hall, Inc., Nueva York, 1946.
(30) Paige, L. J. y Swift, J. D. *Elements of Linear Algebra.* Ginn and Company, Boston, 1961.
(31) Sawyer, W. W. *A Concrete Approach to Abstract Algebra.* W. H. Freeman and Co., San Francisco, 1959.
(32) Schrierer, O. y Sperner, E. *Introduction to Modern Algebra and Matrix Theory.* Chelsea Publishing Company, Nueva York, 1955.
(33) Spain, B. *Analytical Quadrics*, Pergamon Press, Nueva York, 1960.
(34) Sommerville, D. M. Y. *Analytic Geometry of Three Dimensions.* Cambridge University Press, Cambridge, 1934.
(35) Stoll, R. R. *Linear Algebra and Matrix Theory.* McGraw-Hill Book Company, Inc., Nueva York, 1952.

INTEGRACIÓN

(36) Natanson, I. P. *Theory of Functions of a Real Variable*, vol. 2. Traducido al inglés del ruso por Boron, L. L. Frederick Ungar Publishing Company, Nueva York, 1960.
(37) Rogosinski, W. W. *Volume and Integral.* Oliver and Boyd, Edimburgo, 1952.
(38) Saks, S. *Theory of the Integral.* Segunda edición revisada. Traducida del francés al inglés por Young, L. C. Monografie Matematyczne, vol. 7, Warsaw, 1937. Reimpreso por Hafner Publishing Company, Nueva York.

SUCESIONES Y SERIES

(39) Hyslop, J. M. *Infinite Series.* Quinta edición. Oliver and Boyd, Edimburgo, 1954.
(40) Knopp, K. *Theory and Application of Infinite Series*, segunda edición. Traducido al inglés de la segunda edición alemana por Young, R. C., Blackie, Londres, 1951.
(41) Knopp, K. *Infinite Sequences and Series.* Traducido del alemán al inglés por Bagemihl, F., Dover Publications, Inc., Nueva York, 1956.

ECUACIONES DIFERENCIALES

(42) Agnew, R. *Differential Equations*, segunda edición. McGraw-Hill Book Company, Inc., Nueva York, 1960.
(43) Andronow, A. y Chaikin, C. E. *Theory of Oscillations.* Princeton University Press, Princeton, Nueva Jersey, 1949.
(44) Bellman, R. *Stability Theory of Differential Equations.* McGraw-Hill Book Company, Inc., Nueva York, 1960.
(45) Birkhoff, G. y Rota, G.-C. *Ordinary Differential Equations.* Ginn and Company, Boston, 1962.
(46) Burkill, J. C. *The Theory of Ordinary Differential Equations.* Oliver and Boyd, Edimburgo, 1956.
(47) Coddington, E. A. y Levinson, N. *Theory of Ordinary Differential Equations.* McGraw-Hill Book Company, Inc., Nueva York, 1955.
(48) Coddington, E. A. *An Introduction to Ordinary Differential Equations.* Prentice-Hall, Inc., Englewood Cliffs, Nueva Jersey, 1961.
(49) Ford, L. R. *Differential Equations*, segunda edición. McGraw-Hill Book Company, Inc., Nueva York, 1955.

(50) Goursat, E. *Cours d'Analyse*. Séptima edición, vol. II, Gauthier-Villars, París, 1949. Segunda edición francesa del vol. II, parte II, traducida al inglés por Hedrick, E. R. y Dunkel, O., con el título *Differential Equations*. Ginn and Company, Boston, 1917. Reimpreso por Dover Publications, Inc., Nueva York, 1960.
(51) Hale, J. K. *Oscillations in Nonlinear Systems*. McGraw-Hill Book Company, Inc., Nueva York, 1963.
(52) Hildebrand, F. B. *Introduction to Numerical Analysis*. McGraw-Hill Book Company Inc., Nueva York, 1956.
(53) Hildebrand, F. B. *Methods of Applied Mathematics*. Prentice-Hall, Inc., Nueva York, 1952.
(54) Householder, A. S. *Principles of Numerical Analysis*. McGraw-Hill Book Company, Inc., Nueva York, 1953.
(55) Hurewicz, W. *Lectures on Ordinary Differential Equations*. John Wiley and Sons, Inc., Nueva York, 1958.
(56) Ince, E. L. *Ordinary Differential Equations*. Longmans, Green and Co., Londres, 1927. Reimpreso por Dover Publications, Nueva York.
(57) Kamke, E. *Differentialgleichungen. Lösungsmethoden und Lösungen*. Band I, Gewohnliche Differentialgleichungen, 3, aublage, Chelsea Publishing Company, Nueva York, 1948.
(58) Kaplan, W. *Differential Equations*. Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1958.
(59) LaSalle, J. y Lefschetz, S. *Stability by Liapunov's Direct Method*. Academic Press, Inc., Nueva York, 1961.
(60) Lefschetz, S. *Differential Equations: Geometric Theory*. Segunda edición. Interscience Publishers, Inc., Nueva York, 1963.
(61) Leighton, W. *Ordinary Differential Equations*. Wadsworth Publishing Company, Inc., Belmont, California.
(62) Milne, W. E. *Numerical Solution of Differential Equations*. John Wiley and Sons, Inc., Nueva York, 1953.
(63) Minorsky, N. *Nonlinear Oscillations*. D. Van Nostrand Company, Inc., Princeton Nueva Jersey, 1962.
(64) McLachlan, N. W. *Ordinary Non-Linear Differential Equations in Engineering and Physical Sciences*. Oxford University Press, Londres, 1950.
(65) Pontryagin, L. S. *Ordinary Differential Equations*. Traducida del ruso al inglés por Kacinskas, L. y Counts, W. B. Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962.
(66) Scarborough, J. B. *Numerical Mathematical Analysis*, segunda edición. Johns Hopkins Press, Baltimore, Maryland, 1950.
(67) Struble, R. A. *Nonlinear Differential Equations*, McGraw-Hill Book Company, Inc., Nueva York, 1962.

FUNCIONES ESPECIALES Y SERIES DE FOURIER

(68) Carslaw, H. S. *Introduction to the Theory of Fourier Series and Integrals*, tercera edición. Dover Publications, Inc., 1930.
(69) Churchill, R. V. *Fourier Series and Boundary Value Problems*. McGraw-Hill Book Company, Inc., Nueva York, 1941.
(70) Copson, E. T. *An Introduction to the Theory of Functions of a Complex Variable*. Oxford University Press, Londres, 1935.
(71) Jackson, D. *Fourier Series and Orthogonal Polynomials*. The Mathematical Association of America, Buffalo, Nueva York, 1941.
(72) Magnus, W. y Oberhettinger, F. *Formulas and Theorems for the Functions of Mathematical Physics*. Traducido al inglés del alemán por Wermer, J., Chelsea Publishing Company, Nueva York, 1954.
(73) Sneddon, I. N. *Special Functions of Mathematical Physics and Chemistry*. Oliver and Boyd, Edimburgo, 1956.
(74) Rogosinski, W. W. *Fourier Series*, segunda edición. Traducido al inglés del alemán por Cohn, H. y Steinhardt, F., Chelsea Publishing Company, Nueva York, 1959.
(75) Zygmund, A. *Trigonometrical Series*. Dover Publications, Inc., Nueva York, 1955.

Índice analítico